



HAL
open science

Neural correlates of socio-emotional states in macaques

Mina Jazayeri

► **To cite this version:**

Mina Jazayeri. Neural correlates of socio-emotional states in macaques. Neuroscience. Université de Lyon, 2017. English. NNT: 2017LYSE1281 . tel-01768440

HAL Id: tel-01768440

<https://theses.hal.science/tel-01768440>

Submitted on 17 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT :

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de

l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 476

ÉCOLE DOCTORALE NEUROSCIENCES ET COGNITION

Spécialité de doctorat: Neurosciences cognitive

Discipline: Neurophysiologie et neurosciences cognitive

Soutenue publiquement le 18/12/2017, par:

Mina Jazayeri

**Neural correlates of socio-emotional states in
macaques**

Devant le jury composé de:

Mauguière, François, Professeur Émérite, CRNL **Président**

Daprati, Elena, Professeure Associée, Université de Rome **Rapporteuse**

Fadiga, Luciano, Professeur, Université de Ferrara **Rapporteur**

Amiez, Celine, Chercheure, INSERM **Examinatrice**

Andari, Elissar, Chercheure, Université Emory **Examinatrice**

Duhamel, Jean René, Directeur de Recherche CNRS, CNRS **Directeur de thèse**

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	M. le Professeur Frédéric FLEURY
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directrice Générale des Services	Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G.RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y. VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Title: Neuronal correlates of socio-emotional states in macaques

A cornerstone of a successful social life is the ability to correctly predict others' actions and empathically perceive their emotional states. Studies on primates' social interaction have shown that thanks to their keen cognitive abilities monkeys are able to deduce what others can hear or see, and to predict others' emotions and intentions. It has been shown that primates are able to display different degrees of prosocial behavior, from cooperation to even altruism and empathically driven behavior. Studies using fMRI techniques in humans have identified the anterior insula (AI) as a key brain region in the processing of empathy. More precisely, this region emerged as the overlapping area activated for both experienced and observed pain, leading to the idea that empathy for pain may involve a mirror-matching model of the affective and sensory features of others' pain. However, the neuronal basis of this process has yet to be uncovered. In an attempt to extend and to investigate the role of the AI in the process of empathy we have recorded single cell activity in the AI of two monkeys while they were engaged in a social task where based on the performed trials positive or negative reinforcements could be delivered to self, another monkey, or nobody. Behavioral results showed that monkeys take into account the welfare of their partners even when this has no impact on their own welfare. Our neuronal findings report that distinct population of neurons respond differentially to outcomes for self and other, and to appetitive and aversive outcomes. Interestingly the neuronal population responding to the aversive outcome showed mainly three profiles of activity: neuronal representation of conspecifics' unpleasant experience, neuronal representation of own unpleasant experience and a minority of neurons showing mirroring properties between self and other. Thus, our results suggest a neuronal model of empathy that accounts for the distinctive features between feeling and empathizing.

Titre: Les corrélats neuronaux des états sociaux-émotionnels chez le macaque

Un pilier d'une vie sociale fructueuse est la capacité de prédire correctement les actions des autres et de percevoir leurs états émotionnels. Des études d'interaction sociale chez les primates ont montré qu'ils sont capables de déduire ce que les autres peuvent entendre ou voir, et de prédire leurs émotions et intentions. Il a été montré qu'ils peuvent manifester différents degrés de comportements prosociaux, allant de la coopération jusqu'à des comportements altruistes et empathiques. Des études d'imageries fonctionnelles chez l'homme ont identifié l'insula antérieure (AI) comme une région cérébrale clé dans le traitement de l'empathie. Spécifiquement, cette région apparaît comme l'aire intégratrice des activités liées à la douleur ressentie et observée, suggérant que l'empathie pourrait impliquer un modèle « miroir » des propriétés affectives et sensorielles de la douleur d'autrui. Cependant, les bases neuronales de ce processus n'ont pas encore été découvertes. Dans le but d'examiner le rôle de l'AI dans le traitement de l'empathie, nous avons enregistré l'activité des neurones dans l'AI de deux singes pendant qu'ils sont engagés dans une tâche sociale leur permettant de délivrer un stimulus aversif ou appétitif à leur partenaire, à lui-même ou à personne. Les résultats comportementaux ont montré que les singes prennent en compte le bien-être de leur partenaire. Les données neuronales rapportent différentes populations neuronales répondant aux stimuli aversif ou appétitif et ceux délivrés à soi ou à autrui. Notamment, la population neuronale répondant au stimulus aversif a montré trois profils d'activité : une représentation neuronale de l'expérience désagréable du partenaire, une représentation neuronale de sa propre sensation désagréable et une minorité de neurones montrant des propriétés miroirs entre soi et autrui. Nos résultats suggèrent un modèle neuronal de l'empathie représentant des propriétés distinctes entre l'expérience vécue et observée.

Key words: social neuroscience, monkey, empathy, neuronal correlates, anterior insula.

Table of contents

PART I. THEORETICAL PART	6
1. THE PRIMATE SOCIAL NATURE	7
1.1. The social brain	8
1.2. Animal models for “social brain” study.....	9
1.3. Social brain areas	11
2. BEHAVIORAL SOCIAL STUDIES IN NON-HUMAN PRIMATES.....	12
3. NEURONAL CORRELATES OF DECISION MAKING IN A SOCIAL CONTEXT	15
3.1. Human-Monkey interaction studies.....	15
3.2. Monkey-monkey interaction studies with a passive partner	18
3.3. Monkey-Monkey interaction studies with both active agents.....	21
4. EMPATHY.....	25
4.1. The neuroanatomy of empathy: anterior insula and anterior cingulated cortex.....	27
4.2. Animal and human studies of vicarious experience.....	29
4.2.1. Animal studies of vicarious experience	29
4.2.2. Human studies of vicarious experience.....	30

4.3. Electrostimulation in insula in humans and primates.....	32
4.3.1. Electro-stimulation of insula in humans	32
4.3.2. Electro-stimulation of insula in primates	34
4.4. Amygdala: from “body-alarm circuit” to socio-emotional dimension coding and empathy.....	35
4.5. Mirror Neurons.....	38
4.5.1. Mirror system: a circuit for empathy coding?	39
4.5.2. Mirror neurons: sustained and criticized	40
5. PURPOSE OF THE PRESENT STUDY	42
PART II. EXPERIMENTAL SESSION	45
1. MATERIALS AND METHODS.....	46
<i>Animals</i>	46
<i>Animals familiarization with the experimenter and laboratory environment</i>	46
<i>Surgery</i>	47
<i>Head restraint device and recording chamber implantation</i>	48
<i>Behavioral procedures</i>	48
<i>Preliminary training</i>	49
<i>Button pushing and button holding</i>	49
<i>Imperative trials</i>	50
<i>Choice trials</i>	50
<i>Non-social context</i>	51

<i>Social context</i>	51
<i>Imperative protocol</i>	51
<i>Choice protocol</i>	53
<i>Eye tracker system</i>	53
<i>Recording sites definition</i>	54
<i>Electrophysiological recordings</i>	55
2. SCIENTIFIC CONTRIBUTIONS	56
<i>Book chapter: Social Decision-Making in Nonhuman Primates</i>	56
<i>Neuronal basis of empathy in the primate insula.</i>	65
PART III. DISCUSSION	95
The clinical relevance of empathy studies	99
Conclusions and perspectives	101
References	102

PART I

THEORETICAL PART

1. The primate social nature

“Man is by nature a social animal; an individual who is unsocial naturally and not accidentally is either beneath our notice or more than human. Society is something that precedes the individual. Anyone who either cannot lead the common life or is so self-sufficient as not to need to, and therefore does not partake of society, is either a beast or a god.”

Aristotele, Politics

We are social beings and as such we need social contacts. We create social networks through which we feel more realized and complete as individuals in the society. Our efforts for the development of technology over the past years, which allowed us to make connections with others all around the world and opened our minds to new cultures and traditions, represent a clear proof of our sociable nature. It is sufficient to think about the tremendous success of social networks (e.g. Twitter, Facebook, Instagram) which attracted people of any age involved in any field and served as a dynamic tool to show and describe to our conspecifics facts happening in our lives, to share our opinions with others and to stay in contact with them.

Our enthusiastic welcoming of the social media is one of the many examples which showed the most innate nature of the human kind: his will to be “social”. The study of the sociable nature of humans became so attractive for many scientists around the world that a specific domain of neuroscience, “social neuroscience”, has emerged as one of the most popular ones over the past years. Therefore, the research on the brain bases of social cognition and interaction started to move from passive spectator science to studies including engaged participants and simultaneous recordings from the brains of the interacting people (Hari et al., 2015).

1.1. The social brain

Going through human evolution story across centuries a question arises quite spontaneously: what has helped humans to go across all the changes in the society and to survive to them?

Human social life has drastically changed in the course of evolution, we have passed from small realities like life in tribes to bigger ones like life in cities, and as a consequence of this human social behavior has developed as well. The presence of always more complex structured communities forced and challenged the human kind to adapt himself to different situations, and to do this the developing of particularly developed cognitive abilities like theory of mind or recursive thinking were required. The adaptation processes that we have gone through were made possible thanks to our “social brain”. We live in a society where being able to properly communicate with others is the key of our daily life success. The social brain, a set of brain areas dedicated to the elaboration of socially relevant stimuli, is what allows us to correctly predict others actions and empathically perceive their emotional states (Blakemore, 2008). Briefly, it is the social brain that allows us to interact with other people.

Thanks to some of our uniquely human skills, such as language and civilization, we have transformed the planet in a way that no other species has come close to do. Our social brain network allows us to process correctly socially relevant information and thus to properly adapt our behaviors to different circumstances. Currently, many cognitive psychologists and neuroscientists propose two main sets of processes identified in this elaboration: those that are deliberative, controlled and sensitive to context and strategy and those that are automatic and driven by the stimuli (Lieberman, 2007; Adolphs, 2009). These distinctions must be reflected in the neural structures that underlie social cognition and they have become one of the key topics of social neuroscience studies over the past years. To confirm the validity of these studies Firth and Firth (2010) highlighted the central importance of social signals and gave examples of how complex social interactions between two individuals can be investigated under experimental controls in a laboratory (Frith and Frith, 2010).

1.2. Animal models for “social brain” study

To better define the areas of the social brain and to have a deeper understanding of the neural correlates underlying this network, animal models are necessary for experiments that require too invasive techniques to be applied on humans, such as lesion and electrophysiological studies.

For many years, social neuroscience studies have used the rodent model as an animal model, especially focusing on emotional aspects. Recently, there has been a growing evidence that rodents possess a remarkable affective sensitivity to the emotional state of others, which could be developed into experimental models of mental disorders associated with impaired empathy in humans (Langford et al., 2006; Chen et al., 2009; Jeon et al., 2010; Bartal et al., 2011; Burkett et al., 2016). Most of the experiments examined different aspects of rodents' behaviors in response to the distress of conspecifics. Using different experimental paradigms it has been demonstrated that rodents are capable of emotional contagion, specifically for pain (Church, 1959; Langford et al., 2006; Martin et al., 2015) Li et al., 2014; Martin et al., 2015). In the wide literature of observational fear learning studies conducted in rodents it has been shown that social interaction with a distressed partner directly altered the emotional responses of the observers to make a new association (Bredy and Barad, 2008; Guzman et al., 2009; Knapska et al., 2010). Moreover, it has been showed that several factors such as familiarity, the strength of unconditioned stimulus delivered to demonstrator, social interaction, stress, common experience, or a simple sensory cue modulate the degree of rodents' behavioral response to distress in others (Langford et al., 2006; Chen et al., 2009; Jeon et al., 2010; Kim et al., 2010; Atsak et al., 2011; Panksepp and Lahvis, 2011; Sanders et al., 2013; Yusufihsaq and Rosenkranz, 2013; Gonzalez-Liencre et al., 2014; Watanabe, 2015). Other works showed evidences of prosocial and consolation behaviors in rodents. The first robust paradigm to give proofs of prosocial behavior in rats was proposed by Ben-Ami Bartal and colleagues (Bartal et al., 2011). The authors demonstrated that rats learned to release cagemates trapped in a restrainer, even when they did not receive rewards. A more recent study conducted by Burkett and

colleagues (2016) showed that a specific type of rodent species, prairie vole, detects the stress of conspecifics and expresses empathy-based consolation behaviors.

However, the rodent model has its own limitation, especially when thought to be used and compared neuro anatomically to human models. In fact, especially compared to rodents, humans and non-human primates have a hugely elaborated prefrontal cortex (PFC) (Wise, 2008), which is one of the reasons for which non-human primates has established itself as a gold standard to study human's brain.

Indeed, Non Human Primate (NHP), and especially macaque model, has been widely used to study the human social cognition evolution because of its genetic, physiological and neuroanatomical similarities existing between these two species.

It has been shown that compared to other vertebrates, humans and non-human primates have a larger neocortex and this is directly correlated with the size and complexity of their social system (Dunbar and Shultz, 2007). In another study (Sallet et al., 2011), using brain imaging techniques, groups of monkeys have been studied after periods of cohabitation with their conspecifics in social groups of different sizes. Results have revealed that living in larger groups provoked a grey matter expansion in key brain regions for social cognition, such as mid-superior temporal sulcus and rostral prefrontal cortex, and increased coupling of activity in frontal and temporal cortex. Thus, social network size contributes to changes both in brain structure and function.

Therefore, life in society models our brain, so the investigation of human social brain network needs to be conducted in an animal model in which also the social environment dynamics are somehow similar to those of humans. Indeed, primate societies are somehow similar to ours. Many works claimed that the unusual large brain for body size in primates was due to their complex social life (Byrne and Whiten, 1998; Barton and Dunbar 1997; Dunbar, 1998). Even though the tendency for the social brain hypothesis is mostly in terms of group size, the social brain theory has quite explicitly always been about the *complexity* of social relationships (their *quality* rather than merely their *quantity*) (Dunbar and Shultz, 2007). Thus, monkeys live in complex hierarchically structured social environments where, based on the circumstances, they have to compete, cooperate and also show empathically driven behavior with their conspecifics in order to create social bonds and survive.

1.3. Social brain areas

In a first attempt the major components of the social brain have been identified in three main areas: amygdala, orbital frontal cortex and temporal cortex (Brothers, 1990; Machado and Bachevalier, 2006).

The definition of these areas has been possible thanks to both human and non-human primate studies. Primates lesion studies in amygdala have shown how animals become socially isolated after the damage of this limbic brain area (Kling and Brothers, 1992). It has been demonstrated that lesions to monkeys orbitofrontal cortex (OFC) can alter social behavior of the animals which afterwards showed deficits in responding to and producing communicative facial expressions, as well as problems in affiliation and bonding with conspecifics (Raleigh and Steklis, 1981; Babineau et al., 2011; Machado and Bachevalier, 2006). Other findings in primates showed how cells in superior temporal sulcus respond to facial features such as expressions and gaze direction (Perrett et al., 1992).

Whilst all these studies have been conducted in animal models because of their invasive nature, the advent of brain imaging allowed scientists to study the social brain network also in humans. Thanks to human studies there have been two major additions to the first list of social brain regions defined by Brothers: first, the medial prefrontal cortex (mPFC) and the adjacent paracingulate cortex which have been implicated in studies where participants were asked to think about mental states (Amodio and Frith, 2006); second, a “mirror system” in ventral premotor cortex that has been found in both humans and monkeys’ brain which allows us to share the experience of others as if we are experiencing them ourselves just by mere observation (Di Pellegrino et al., 1992; Rizzolatti and Craighero, 2004).

Thus, despite the evolutionary expansion of the human brain and the increased complexity of our social networks, the findings in primate studies described above highlight that the brain regions involved in social cognition in humans and macaques are impressively similar. For this reason, the social nature of primates has been investigated from both a behavioral and neuronal point of view in several studies that are going to be described in the following chapter.

2. Behavioral social studies in non-human primates

Several studies have been conducted in macaques while they were inserted in social contexts to evaluate their behavior.

Studies on primate social interaction have shown that monkeys are able to deduce what others can hear or see (Rushworth et al., 2013; Flombaum and Santos, 2005; Santos et al., 2006), or to predict others' emotions and intentions (Cheney et al., 1986). Early experimental studies have shown that macaques are able to show "altruistic" behavior by sacrificing personal interests in order to alleviate their peers' distress (MASSERMAN et al., 1964; Miller et al., 1966). Other studies showed how monkeys take into account the welfare of their conspecifics by using a simple decision making task which involved choosing between rewarding just self or self and a passive partner at the same time (Massen et al., 2010). In a more recent study, by using a challenging social decision making task involving both appetitive and aversive stimuli, Ballesta and colleagues (2015) showed how monkeys can behave prosocially by choosing to reward a conspecific instead of sending the juice to a non-living agent and by preventing the partner from getting the aversive outcome, which consisted in an airpuff, by sending it to a non-living agent. Furthermore, thanks to physiological measurements such as the eye blink rate of the animals while they were getting the airpuff, they showed how monkeys react empathically by observing their conspecifics getting an airpuff as if they were experiencing themselves the unpleasant sensation felt from the other. All these elements allowed the authors to talk about empathy driven behaviors in macaques (Ballesta and Duhamel, 2015). The work of Ballesta and colleagues used a very challenging experimental paradigm for non-human primates. In their task, animals were asked to determine not only their own faith but also that concerning their partners. Actor monkeys had the free will to choose what was going to happen to their partner while they were controlling the task, and they chose the prosocial option most of the times. But what allows the authors to go further than the pro-social behavior of the animals is the fact that monkeys were surprisingly *reacting* to what was happening to their partners as if they were somehow "feeling" it themselves.

All these findings obtained from the behavioral studies described above show how primates can behave in a prosocial manner by taking into account the welfare of their conspecifics.

Since it has been shown that monkeys are able to make social choices, several studies started to combine decision making tasks with electrophysiological recordings to investigate the neural correlates of primate social behavior. This network includes the dopamine reward system, amygdala complex, temporal cortex, insula, the anterior cingulate (ACC), ventromedial (VM) and orbitofrontal (OFC) subdivisions of the prefrontal cortex (Figure 1).

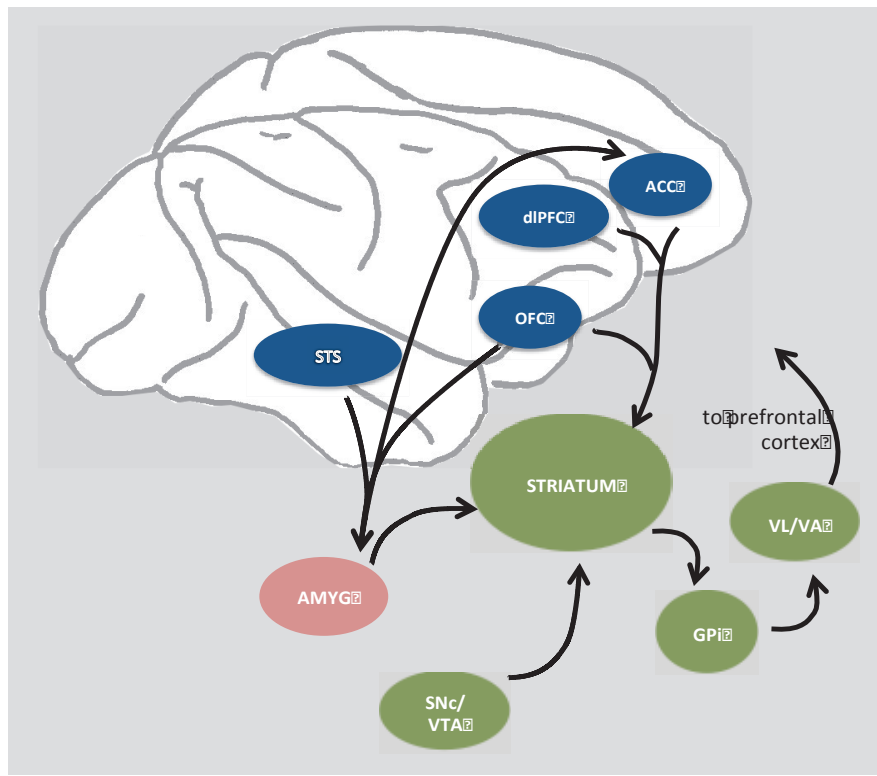


Figure 1. Simplified schematic representation of the primate reward and social decision-making network. Midbrain nuclei containing dopaminergic neurons: SNc/VTA, substantia nigra pars compacta / ventral tegmental area. AMYG, amygdala; GPI, internal globus pallidus; VL/VA, ventral lateral and anterior thalamic nuclei. Subdivisions of the prefrontal cortex: STS, superior temporal sulcus, OFC, orbitofrontal cortex; dIPFC, dorsolateral prefrontal cortex; vmPFC, ventromedial prefrontal cortex, ACC, anterior cingulate cortex (Font: Chapter 14, Decision Neuroscience: An Integrative Perspective).

3. Neuronal correlates of decision making in a social context

3.1. Human-Monkey interaction studies

Non-human primates are smart animals capable to learn by observing their conspecifics' actions and this has been widely observed in ethological studies.

Several studies have explored this capacity using token exchange paradigms where tokens are meant to be objects with a symbolic meaning and no intrinsic value. Objects can become tokens when they acquire some value through arbitrary associations made with what is returned in exchange. Studies conducted in capuchins (Brosnan et al., 2004a; Brosnan et al., 2004b; Westergaard et al., 1998; Addessi et al., 2007) (Falcone et al., 2012) and chimpanzees (Brosnan et al., 2005) have examined the ability of these animals to exchange tokens for food or tools. Besides methodological differences all these works have shown that not only the symbolic meaning of a token can be associated with a reward but animals can also associate different quantities of rewards with the corresponding tokens.

Token exchange paradigms have also been examined using observation learning procedures.

Learning by observation is an adaptive ability present in humans, both adults and children, non-human primates and many other species such as mice (Zentall et al., 1972), fishes (Schuster et al., 2006), reptiles (Davis et al., 2011) and insects (Leadbetter et al., 2009). Most of the time, learning by observation helps humans and animals to learn faster avoiding the learning process through error trials which are often necessary, especially in the process of learning something *ex-novo*.

By using a token exchange paradigm, Brosnan and de Waal (2004) showed the learning of tokens' value through observation of conspecifics' actions. They showed how this observation-learning process was limited by the fact that the observed actions were performed by conspecifics (Brosnan et al., 2004b).

However, a study of Falcone and colleagues (2012) showed how monkeys were actually able to learn by observing also non conspecifics' actions, specifically humans (Falcone et al., 2012b). In their study they presented the monkeys with a human model who demonstrated them how to solve an object-reward association dilemma to obtain some food reward. Human models were presented with a pair of neutral objects where they had to choose the correct one in order to obtain the food reward associated with it. This phase of the study was considered as a learning phase where monkeys could learn by observing humans' actions. In the following test phase the monkeys made a choice of their own. The performance of the animals in the test phase confirmed the ability of monkeys to learn by vicarious observation of human models.

In another study using a token exchange paradigm it has been shown how monkeys were able to learn from humans' actions by observing human subjects exchanging three different objects with the experimenter among which only one was associated with a reward (Bevacqua et al., 2013). This learning by observation phase was followed by a test one where monkeys had to perform the task on their own with the particularity that they had to choose directly the previously chosen object by the human among the three presented to them in order to obtain the reward. The results showed again the ability of the monkeys to learn from a non-conspecific actor through the experience of the vicarious reward.

But monkeys are able to monitor also more complex actions performed by humans and even to successfully and cleverly collaborate with them, especially if by doing so they can achieve a reward for themselves. Falcone and colleagues (2012) conducted a study by using a non-match-to-goal task which involved human-monkeys interactions. In this paradigm, they had to follow the precise rule to reject the previously chosen goal in order to correctly perform the task and obtain a reward (Falcone et al., 2012a). During the task, monkeys were presented with a pair of targets from a list of three and they had to choose one target in the presented pair. Then, in the next trial they had to avoid making the previous choice switching to the alternative one. In a subset of trials monkeys performed the same task but in collaboration with a human partner. Once the human partner concluded his turn, the monkeys had to takeover and continue the task and to do this correctly they had to switch to a new goal discarding the human's previous goal. The findings showed how monkeys were

monitoring not only their own choices but also those of their human partner and highlighted the impressive ability of the animals to coordinate their actions with those of a non-conspecific.

All these human-monkey interaction findings suggest the ability of primates to monitor the action of an agent who can also be a non-conspecific. This led to the hypothesis that the sense of agency is something cognitively well conceived from monkeys.

Actions are normally associated with goals, independently from who is performing them. We do things to achieve specific goals that might be crucial for our survival. One of the brain regions that represents goals, both past and future, is the primate prefrontal cortex. Falcone and colleagues (2015) investigated the role of the prefrontal cortex in monitoring other agent's goals by using a non match-to-goal task where monkeys and human partners switched actor and observer roles (Falcone et al., 2015). The task design that they used was very similar to the one used in their previous study (Falcone et al., 2012), despite some elements that made the performing of the task even more challenging for the animals, like the increase of the number of targets. Their results confirmed once again the ability of the animals to perform the task by successfully cooperating with their human partners. From a neuronal point of view, they found neurons in lateral prefrontal cortex (LPFC) encoding the actor, either the monkey itself or its human partner, neurons encoding the agent's future goal position and neurons encoding the agent previous goal position. Interestingly, the subpopulation of cells involved in the computation of the human future goal showed mainly two profiles of activity: some of these cells were encoding the future goal of both agents while others were encoding exclusively the human agent future goal. Briefly, this brain region does encode a social dimension in terms of agency and therefore is one of the crucial elements of the social cognition network.

All the studies described till now have focused on the investigation of the social nature of primates. Some of them have evaluated if monkeys were able to conceive a social dimension by observing others' actions, thus by elaborating the sense of agency.

3.2. Monkey-monkey interaction studies with a passive partner

Several electrophysiological experiments have investigated the neuronal correlates of motivated behaviors and reward processing in non-human primates, especially in the dopamine reward system (Schultz, 2015).

The representation of different natural reinforcers, both appetitive (e.g. juice) (Rolls et al., 1994; Rolls, 2000) and aversive (e.g. airpuff) (Morrison and Salzman, 2009), has been described in the orbitofrontal cortex (OFC), and it has been demonstrated that activity in this brain area reflects the subjective value of rewards during decision making (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006).

In an attempt to extend and to investigate the role of the OFC from subjective value-based decisions to social ones, Azzi et al. (2012) trained monkeys to perform a task where animals had to perform manual responses to visual cues predicting juice rewards for themselves only and for themselves and passive partners simultaneously (Azzi et al., 2012). In this study, animals were performing the task inserted in a social context, meaning that they were always in presence of their passive conspecifics while performing the trials. Neuronal results showed mainly two profiles of activity of cells recorded in this area: a) OFC neurons responded to the subjective value of rewards with a higher discharge for big amount of rewards compared to smaller ones, and this was perfectly coherent with the behavior of the monkeys who showed a better performance in trials predicting large amount of rewards compared to those predicting small amount of juice; b) neuronal activity was modulated by the social context always in a coherent way with the behavior of the animals. Actor monkeys showed lower performance in the trials where they were rewarded simultaneously with their passive partners compared to those where they were the only rewarded ones, and at a neuronal level this was reflected in a lower spontaneous spike discharge activity of the cells in the OFC when the actor monkey was consuming the reward in parallel with its passive partner compared to trials where he was consuming the reward alone. Other results from this study suggested that monkeys did also have social preferences for some partners instead of others since they were more motivated performing trials where they were rewarded at the same time with the preferred partner rather than

those where they were working for themselves and the non-preferred one. This social preference was reflected at a neuronal level with cells showing higher firing rate when the actor monkey was working for himself and the preferred partner, but lower activity when the same monkey was providing reward to himself and the non-preferred conspecific. Such properties have shown how the OFC could be involved in the neural computations taking place in social contexts and how its activity can be modulated by these environments.

However, thinking about human nature it is clear how much we do have the tendency to compare our own welfare with the one of others when we are in a social context and especially in presence of our conspecifics. We often compare our own goods with the ones of other individuals thinking about equity and inequity between what we have and what others have. The study of Azzi et al. (2012) does not address directly the coding of subjective equity and inequity between own and other's welfare.

It has been already shown that non-human primates respond adversely to social inequity (Brosnan, 2013; Proctor et al., 2013). Human studies in adults and children have shown that the striatum is active in relation to other's reward, to reward inequity and also during the learning process about social agents (Adams et al., 1965; LoBue et al., 2011).

The work of Báez-Mendoza and colleagues (2016) investigated specifically neuronal coding of inequity in the ventral striatum of primates (Báez-Mendoza et al., 2016). In their experiment they used an imperative reward-giving task in which monkeys faced four different conditions: reward to self, reward to the partner, reward to both animals or reward to nobody. Behaviorally, animals showed a moderate inequity sensitivity in the laboratory setting. At a single cell level, as shown in prior studies (Hikosaka et al., 1989; Cromwell, 2003), a sizeable proportion of striatal neurons encoded reward to self. Interestingly, other subpopulations of neurons in this brain region coded either disadvantageous inequity (receiving less reward than the conspecific) or advantageous inequity (getting more reward than the conspecific), while few striatal neurons coded both inequity forms. Neurons in the striatum can also respond to another individual's actions, but only when these actions are linked to own reward (Baez-Mendoza et al., 2013).

Overall, the picture that emerges from the several works conducted in the striatum in

humans and non-human primates is that this subcortical region is a key brain area capable of integrating social information into coding of social actions related to own rewards.

So far the described works have focused on brain regions involved in the computation of subjective utility of own rewards as a function of the social context. Thus the “frame of reference” of both the OFC and the striatum remains self centered and does not purely describe a “not-self” dimension, which could be a live or a non-live agent, experiencing the receiving of a reward or performing an action that does not necessarily lead to a self reward.

In an attempt to see whether the “non-self” dimension could be computed in frontal cortical areas such as the OFC, the anterior cingulate gyrus (ACCg) and the anterior cingulate sulcus (ACCs) Chang and colleagues (2012) recorded neurons in these brain regions while animals were performing a reward allocation task in the presence of a passive conspecific. Actor monkeys were presented with three main conditions: reward to self, reward to the partner and reward to a non live agent (neither) (Chang et al., 2012). These conditions were then combined to make the animals perform choice trials giving them the freedom to cognitively choose the recipient of the reward. Their behavioral results showed how monkeys behaved in a prosocial way by choosing to reward the partner instead of nobody. From a neuronal point of view their findings showed that neurons in the in the OFC selectively encode own rewards and this is consistent with previous studies implicating this area in representing the subjective value of rewards (Tremblay and Schultz, 1999; Padoa-Schioppa et al., 2006). Neurons in the ACCs encoded reward allocations to the partner monkey or no one, thus the foregone rewards. These results on the ACCs neurons activity were consistent with previous studies showing the implication of this brain region in error monitoring and behavioral adjustment (Carter et al., 1998; Alexander et al., 2011).

Thus, within this network of received and forgone reward signaling respectively in the OFC and ACCs, ACCg emerges as a key brain region for the computation of shared experiences and social reward. Indeed, this “mirroring” of self and other rewards by ACCg neurons is perfectly consistent with previous works showing properties of this area in encoding social variables such as shared experiences and empathy (Amodio and Frith, 2006).

In all the studies described above there is one common factor in the experimental paradigms that the authors used: in all of them just one of the monkeys was actively performing the task while its partner was completely passive. Clearly in contexts like this the possibility to study neural correlates of the sense of agency related to the *other* is impossible since the other is not doing anything but just passively waiting that something is chosen or not for him by the actor monkey, thus the only sense of agency that can be studied is self related. However, it is well known how we normally monitor other's actions, independently from if they have consequences or not for us. We commonly call it *curiosity* which is typical in children and adults and is crucial in the learning process. To study the neuronal correlates related to the observation and to the monitoring of other's actions some experimental paradigms involving dynamic interactions between pair of monkeys were developed (in these studies both monkeys were actively involved in the task by performing in an equal measure the trials). These works are described in the following chapter.

3.3. Monkey-Monkey interaction studies with both active agents

One of the brain areas involved in the representation of other's actions is the medial frontal cortex (MFC) (Yoshida et al., 2011). In their study, Yoshida and colleagues trained a pair of monkeys to do a role reversal task in which monkeys sat face to face and took turns making a choice to obtain a reward. In this kind of behavioral task monkeys could monitor not only their own actions but also the ones that were being performed by their partners. Their findings show that neurons in the MFC encode self and other's actions and distinguish self and other's dimension in the motor domain. This way of computation of the social dimension in the MFC, which is in terms of agency, shows how this brain region is crucial for social learning.

Although much learning comes from observing others' actions it is a clear fact that most of the times we learn from our mistakes. We are human beings and despite our cognitively sophisticated way of thinking and reasoning about our daily life experiences we do make mistakes. As Oscar Wilde says: "experience is simply the name we give our mistakes." Most of the times by making mistakes we are, in a first attempt, sorry about them but *a*

posteriori we use them to learn. The same mechanism can occur in a social context. We can learn from others' mistakes just by observing them.

This mechanism of learning from others' mistakes must have neural correlates underlying it, and in their work Yoshida and colleagues (2012) investigated these by using a social interaction task in monkeys (Yoshida et al., 2012). The medial frontal cortex (MFC) has previously been shown to process self-generated errors (Niki et al., 1979 2; Ridderinkhof et al., 2004 5). Moreover, several human studies using event related potentials and functional neuroimaging have shown that MFC is associated with others' error monitoring processes (Van Schie et al., 2004 6; Miltner et al., 2004 7; Shane et al., 2004 9). In their study, Yoshida et al. addressed this issue by investigating the activity of individual neurons in the MFC using the same role reversal task they used in their previous work (Yoshida et al., 2011). By performing the trials to obtain a reward, monkeys occasionally made mistakes, and these errors were observed by the partner who was waiting his own turn. They identified a group of neurons in the MFC which encoded the other's errors. Partner's error could be of two types: a) errors leading the monkey not to accomplish the trials correctly, having as consequence the reward omission and b) other's erroneous actions. Nearly half of the neurons showed activity modulation consistent with reward omission signals, whilst the remaining neurons responded to other's erroneous actions. Their findings showed that the MFC is not only involved in the coding of social agency but also in the monitoring of other's mistakes during social interaction, thus confirming its important rule in the learning process.

In both animal and human's society some resources for survival and reproduction in nature are limited, thus competing successfully with other individuals to obtain such vital elements turns out to be crucial. The course of history has shown that humans are able to successfully and strategically compete with their conspecifics, conquering countries and building up empires. As humans, sometimes we even enjoy competitions and the fact itself to challenge ourselves with other individuals, and we like this even more when we are the winners. In a fascinating way this aspect of humans' society can be seen in primates' reality. Indeed, primates live in highly social environments characterized by strong competition and dominance hierarchies where they have to strategically compete with their

conspecifics for vital resources such as food and sex. Thus, understanding the neuronal correlates of competitive behavior is crucial for the study of the human being in the social context, and because of the similarities between human and non-human primates this species could be used as a relevant model.

Hosokawa and Watanabe (2012) studied individual neurons in lateral prefrontal cortex (IPFC) of monkeys while the animals were performing a competitive task. In the behavioral task monkeys were playing a video shooting game in three different contexts: competing against a conspecific, playing against a computer or playing alone without a rival (Hosokawa and Watanabe, 2012). Behaviorally, they observed that animals were performing better and faster in the competitive contexts, thus when they were faced with a competitor which could be a live (a conspecific) or a non-live (a computer) one, compared to the non-competitive context in which they were performing the task alone. This behavioral differentiation was reflected in the neuronal activity of the IPFC cells. Indeed, neurons in this region responded differently between the competitive and non-competitive games showing also winning-losing related activity. Moreover, activity of cells in this region differed depending on whether the competition was between monkeys or between the monkey and the computer. All these findings together suggest that the IPFC is involved in the computation of the concept of competition and encodes the result of the competition as well. Thus, the correct functionality of this brain region is crucial to survive in a complex and competitive social environment since it helps animals to correctly and advantageously adapt their behavior to the circumstances in a way that might increase their chances to win a competition and obtain rewards.

All the studies described above highlight the ability of primates to detect others' "known" actions, where known means observable, thus the coding of agency as self and other when self and other's actions are known and observed. Other prior studies have described a brain network, the so-called "mirror system" in which single neurons encode another's known and observed actions as well as own performed actions (di Pellegrino et al., 1992; Rizzolatti and Sinigaglia, 2010). But whether and what neurons encode other's unknown actions is something that hasn't been shown in the described works above. Another's imminent intentions are practically unobservable, yet there must be a neuronal coding of these

actions. The existence of these neuronal population has been hypothesized by several works focusing on animal social behavior (Firth and Firth, 1999; Gallese and Goldman, 1998; Rilling et al., 2004; Sanfey et al., 2006; Vogeley et al., 2001), but it has never been proved till the study conducted by Haroush and Williams (Haroush and Williams, 2015). The authors conducted a study by using a joint decision paradigm to study mutual decisions in primates and explored neuronal evidences that predict another agent's intentions and ways of cooperation. In their behavioral task monkeys were sitting side by side facing a screen and they could not see the reciprocal decisions they were making in real time, thus the actions of the other monkey were unknown to them. The choice terms, cooperation and defection, were derived from Prisoner's Dilemma (iPD) literature (Camerer, 2003). On each trial animals covertly chose between two options and only after a delay their choices were revealed to the partner and the reward associated to them was delivered to the animals. Authors specifically focused on the study of cells' activity in the dorsal portion of the anterior cingulate cortex (dACC), a regions strongly connected with frontal and temporal-parietal areas which have been shown to be involved in interactive behavior (Behrens et al., 2009; Paus, 2001) to its role in encoding social interest in other individuals based on functional imaging (Behrens et al., 2008) and to ablative studies conducted in there (Rudebeck et al., 2006b). The results showed that dACC neurons encoded the monkey's own decision to cooperate while other subpopulation of cells predicted exclusively another agent's yet unknown decisions during social interaction. Another question has been whether monkeys are able to take into account what is happening to their conspecifics as a consequence of something that they do for them or their conspecifics do for themselves. This issue has been addressed in several studies which used social decision making tasks and showed how monkeys mostly chose the advantageous options for their partners (Chang et al., 2012; Ballesta et al., 2015). In their work Ballesta and Duhamel (2015) went further than the demonstration of the prosocial nature of primates and showed evidences of empathy in macaques. Thus, macaques are able to take into account the welfare of their conspecifics and also to react to what is happening to them as if they are involved in first person in the sharing of their emotional states, be them positive or negative.

Such behavioral evidences of empathy in primates must have neural correlates underlying them. In the next chapter we will first go over the general definition of empathy and some

of the theories that talk about the structures of empathy with an overview on all the scientific experiments conducted on humans and non-human primates to investigate which brain regions are involved in the computation of this feeling.

4. Empathy

“A prerequisite to empathy is simply paying attention to the person in pain.”

Daniel Goleman

The word “empathy” was first introduced by the psychologist Edward Titchener over 100 years ago as a translation of the German word *Einfühlung* (“feeling into”). There are many definitions of empathy and none of them has been accepted as the universal one, but overall most theories agree that empathy is the ability to vicariously experience and to understand the feeling of other people (Eisenberg et al., 1990; Hoffmann et al., 2008; Singer et al., 2009; Bird et al., 2014).

Within the structure of empathy two dimensions have been distinguished: the emotional/affective and the cognitive ones.

The affective form of empathy is commonly referred to as an affective state (such as the experience of a pleasant or unpleasant emotion) caused by sharing the state of another person through observation or imagination of their experience (i.e. the vicarious experience) (de Vignemont and Singer, 2006; Singer and Lamm, 2009). Although an observer’s emotional state is isomorphic with that of another individual, the observer is aware that someone else is the source of that state (de Vignemont and Singer, 2006). By contrast the cognitive aspects of empathy are commonly associated with theories of perspective taking, mentalizing or theory of mind (Frith and Frith, 2006).

Perspective taking skills are rooted in a cognitive skill called “Theory of Mind” (ToM). Perspective taking could be defined as the capacity to take others’ affective perspective: for example understanding their specific situations and needs, separate from own ones, which still requires access to personal representations of the other’s state. The ability to “mentalize” is to understand and manipulate other people’s behavior in terms of their

mental states, which is a major ingredient in successful social interactions (Singer, 2006)(Frith et al., 1991, 1999). Premack and Woodruff (1978) defined the capacity to understand that other people have beliefs and desires different from our own as “theory of mind”.

It is important to note that some authors define empathy with only its affective components and consider the cognitive dimension as a separate but related construct of “theory of mind” or “mentalizing” on the basis that these processes rely on large distinct neurocognitive circuits (Singer, 2006). Even intuitively people associate empathy with something linked to an emotional sphere rather than something linked to a cognitive one. But it is important to underline that the affective state of empathy could be something experienced even by babies and toddlers till the age where their rational way of thinking is sufficiently developed to allow them, for instance, to distinguish self from other. Indeed, without this distinction it is quite hard to see things from the perspective of another person. We need to know who the *other* is before seeing things from his/her perspective, or in other words to “put ourselves in someone else’s shoes”. This step is crucial for the process of empathy and is mediated by its cognitive components. Thus, combined together, these dimensions of empathy enable us to understand another person’s beliefs, desires and emotions (Firth and Firth, 2006).

However, affective empathy is well distinguished from concepts such as emotion contagion, mimicry, empathic concern, compassion and sympathy (Singer and Lamm, 2009; Bird and Viding, 2014). These processes risk to be confused with the concept of empathy since they usually occur in similar contexts, thus separating them is crucial and not easy at all. For example, a recent model of empathy called the self-to-other model of empathy (SOME) (Bird and Viding, 2014) highlights that emotional contagion is a key precursor to empathy but does not have to involve a distinction between self and other. Thus, although emotional contagion may be necessary for empathy, and is an instance of a vicarious experience, on its own it is not sufficient due to a lack of self-other distinction. Empathic concern, which is commonly known as “sympathy” or “compassion”, involves having feeling for another individual and is associated with motivation to alleviate their

suffering. However since empathic concern does not necessarily involve any vicarious experience it is well distinguishable from affective empathy.

Many studies have focused on the behavioral and neural correlates of vicarious experience in humans and monkeys and new data allowed the development of empathy models such as the perception-action model (PAM) and mirror neuron theories. Although, before describing these works it is important to have an idea of the neuroanatomy of empathy.

4.1. The neuroanatomy of empathy: anterior insula and anterior cingulate cortex

The anterior cingulate cortex (ACC) and anterior insula (AI) are key brain regions that respond during vicarious experiences (Bernhardt and Singer, 2002; Lamm and Decety, 2011, Lockwood et al., 2015). Thus, understanding their functional anatomy is crucial to understand how vicarious information is processed in the brain (Figure 2).

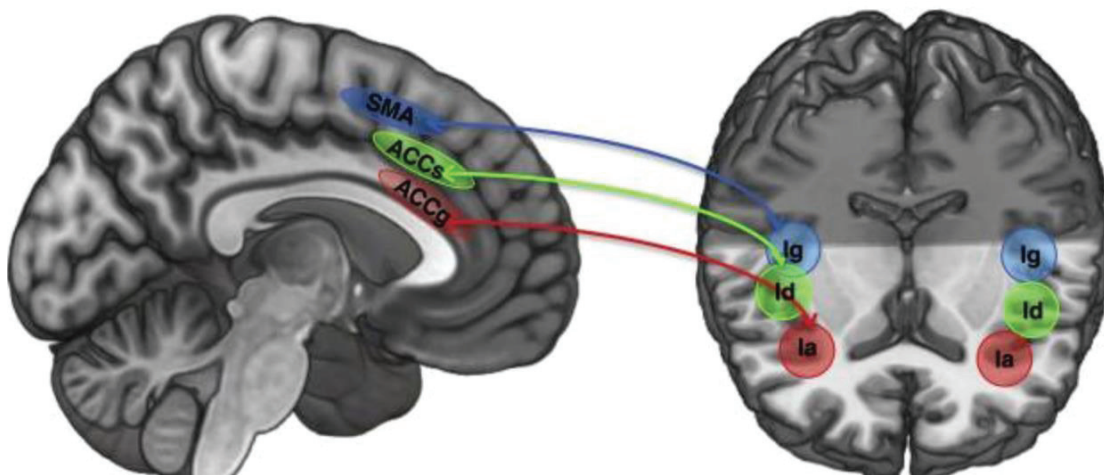


Figure 2. Connectivity between cytoarchitectonic sub-regions of the cingulate cortex and insula. SMA = supplemental motor area, ACCs = sulcal portion of the anterior cingulate cortex, ACCg = gyral portion of the anterior cingulate cortex, Ia = agranular anterior insula, Id = dysgranular mid insula, Ig = granular posterior insula (Font: Lockwood, 2016).

The cingulate cortex is anatomically and functionally heterogeneous and comprises distinct cytoarchitectonic zones. These areas have been labelled as retrosplenial, posterior, medial (MCC) and anterior cingulate cortex (ACC) (Palomero-Gallagher et al., 2008). Both MCC and ACC are subdivided in the sulcus (MCCs/ACCs, henceforth ACCs) and in the gyrus (MCCg/ACCg henceforth ACCg) that have different functional properties. The ACCs has connections to primary motor, premotor, supplementary motor (SMA) and pre-supplementary motor (pre-SMA) cortices intraparietal sulcus, orbitofrontal cortex and nucleus accumbens (Showers, 1959; Wang et al., 2001). Posterior portions of the ACCs are often considered motor areas based on their direct projections to the spinal cord (Hutchins et al., 1988; Dum and Strick, 1996). Indeed electrical stimulation of neurons in ACCs induces limb movement (Luppino et al., 1991). The ACCg has connections to posterior portions of the superior temporal sulcus, the temporo-parietal junction and dorsomedial prefrontal cortex (dmPFC) (Pandya et al., 1981; Seltzer and Pandya, 1989; Barbas and Ghashghaei, 1999) that are known to be involved in the processing of the mental states of others (Firth and Firth, 2006). Importantly, the ACCg has strong connections to anterior but not posterior insula (Mesulam and Mufson, 1982). Both ACCs and ACCg have common connections to medial and lateral portions of the orbitofrontal cortex (Morecraft et al., 1992; Morecraft et al., 1998) and to the nucleus accumbens (Haber et al., 1995; Kunishio et al., 1994) suggesting the involvement of both regions in processing rewards.

The insula is also an anatomically and functionally heterogeneous brain area. Based on the degree of granularity, modern descriptions of the insula generally agree on three subdivisions which are anterior agranular cortex (anterior insula), a middle dysgranular cortex (middle insula) and a posterior granular cortex (posterior insula) (Mesulam and Mufson, 1982). It has been shown that these subregions have distinct connectivity patterns in both human and non-human primates (Mesulam and Mufson, 1982; Taylor et al., 2009; Shura et al., 2014). The AI has connections to the ACCg, frontal operculum, OFC, dorsal and ventral temporal pole, and sensory areas such as the somatosensory and opercular areas of the parietal lobe (Mesulam and Mufson, 1982; Taylor et al., 2009; Shura et al., 2014). The middle insula has connections to the ACCs, frontal operculum, ventromedial prefrontal cortex (VMPC), OFC, to the secondary somatosensory area, to the superior temporal sulcus, ventral striatum and amygdala (Mesulam and Mufson, 1982; Taylor et al., 2009;

Shura et al., 2014). The posterior insula is connected to the SMA, VMPC, temporal poles, secondary somatosensory area, and dorsolateral striatum (Mesulam and Mufson, 1982; Taylor et al., 2009; Shura et al., 2014). The posterior insula receives projections from the spinothalamic pathway, the major pathway for processing nociceptive information, whereas these projections do not seem to reach AI (Dum et al., 2009). Stimulation of neurons in the posterior insula elicits feelings of pain and warmth, and this hasn't been observed stimulating other regions of insula (Shura et al., 2014). Importantly, the AI connects to the ACCg whereas the mid and posterior insula are primarily connected to ACCs and SMA respectively as shown in Figure 2.

Overall, the anatomical and functional profile of ACCg and AI suggest that these regions may be involved in processing social information, more specifically information that is directed to or about other people; and that's why many studies of vicarious experience in humans and animals have explored these brain areas.

4.2. Animal and human studies of vicarious experience

4.2.1. Animal studies of vicarious experience

Several studies have been conducted in humans and animals to see which brain regions are involved in the elaboration of vicarious experience, and among these brain regions the anterior insula (AI) and the anterior cingulate cortex (ACC) have been those linked directly to the process of empathy.

Research in non-human primates has mostly focused on the observation of others' rewards rather than pain, and on the ACC rather than the AI. There are evidences that the ACCg plays a key role in social cognition and behavior in both humans and NHPs (Lockwood et al., 2015; Apps and Green, 2013; Sallet et al., 2011). In particular animal models have suggested that there are important divisions between the ACCg and ACCs that are crucial for understanding social behavior (Chang et al., 2013; Rudebeck et al., 2006). In their work Lockwood and colleagues (2013) argue that while both the ACCg and ACCs are involved in the processing information that conforms to the principles of the reinforcement learning

theory only the ACCg does this in social contexts while the ACCs does so in “non-social” ones (Lockwood et al., 2013).

Moreover, lesion studies in ACCg showed an impairment in the processing of social stimuli and a deficit in the execution of social behaviors, whereas lesions to ACCs and OFC did not (Rudebeck et al., 2006).

Finally, in their work Chang and colleagues (2013) showed how the ACCg shows mirror-like activities in a social context where self and other’s rewards were encoded in a similar way suggesting the involvement of this brain region in the processing of empathy (Chang et al., 2013).

4.2.2. Human studies of vicarious experience

While most of the animal studies of vicarious experience have focused on the observation of others reward rather than pain and on the anterior cingulate cortex (Chang et al., 2013; Rudebeck et al., 2006; Lockwood et al., 2015), many human studies of vicarious experience have focused on the observation of other people in pain.

One of the first studies to investigate the neuronal responses to the observation of other people’s pain was conducted by Singer and colleagues (2004) using functional magnetic resonance imaging (fMRI) on subjects experiencing a painful stimulus while undergoing a scan (Singer, 2004). Subjects were then presented with cues signaling that their partner, present in the same room, was going to receive a painful stimulus too. They showed that anterior insula and anterior cingulate cortex responded both in the self and other’s trials.

Since this finding other fMRI studies have been conducted on humans by using cue-based paradigms, which showed to the subject cues that had painful consequences for the other (Singer et al., 2004; Jackson et al., 2005; Hein et al., 2010), or pictured based paradigms where subjects were presented with pictures of body parts likely to be hurt (Jackson et al., 2005; Lamm et al., 2007). All together these findings support the idea that the direct experience of pain and the observation of others painful experiences activates similar neural

regions, especially the anterior insula, which is thought to be a neural marker of empathy (Figure 3 shows the neural network underlying empathy for pain).

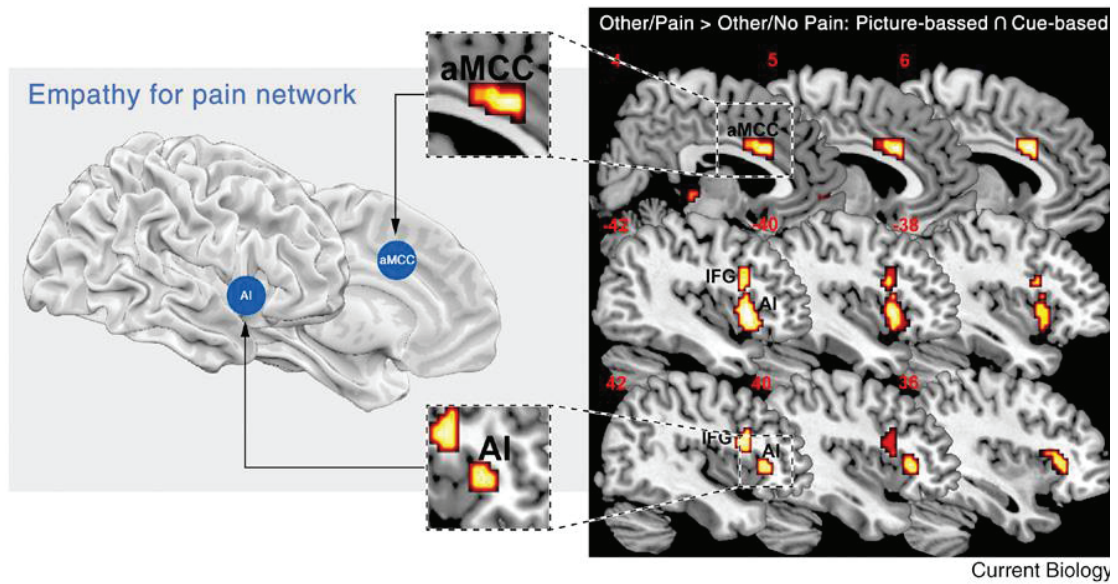


Figure 3. Neural network underlying empathy for pain. Depicted functional neural activations on the right are the result of a meta-analysis based on nine fMRI studies investigating empathy for pain. AI, anterior insula; aMCC, anterior middle cingulate cortex; IFG, inferior frontal cortex. (Font: Singer et al., 2014).

4.3. Electrostimulation in insula in humans and primates

4.3.1. Electro-stimulation of insula in humans

Numerous functional imaging studies conducted in humans confirmed the involvement of the insular cortex in processing painful as well as non-painful somatosensory inputs shown in primates anatomical and microelectrode studies (Burton et al., 1993; Casey et al., 1994, 1996, 2001; Coghill et al., 1994; Hsieh et al., 1995b; Craig et al., 1996, 2000; Vogt et al., 1996; Andersson et al., 1997; Antognini et al., 1997; Derbyshire et al., 1997; Rainville et al., 1997; Svensson et al., 1997; 1998; Xu et al., 1997; Davis et al., 1998a,b; Derbyshire and Jones, 1998; disbrow et al., 1998; Iadarola et al., 1998; May et al., 1998; Oshiro et al., 1998; Paulson et al., 1998; Bushnell et al., 1999; Gelnar et al., 1999; Ploghaus et al., 1999; Kwan et al., 2000). Laser evoked potentials in humans have shown bilateral dipolar sources in the second somatosensory area (SII) or the insular cortex (Tarkka and Treede, 1993; Bromm and Chen 1995; kakigi et al., 1995; Valeriani et al., 2000; Opspmmer et al., 2001).

However, the first study on direct electrical stimulation of human insular cortex was conducted by Ostrowsky and colleagues (Ostrowsky et al., 2002). Direct stimulation of the insular cortex during the pre-surgical phase of epileptic patients has always been a challenge because of its anatomic location in the brain. Indeed insular cortex is buried under the frontal, temporal and parietal opercular cortices and covered by a dense wall of vessels. A representation of human insula is shown in Figure 4.

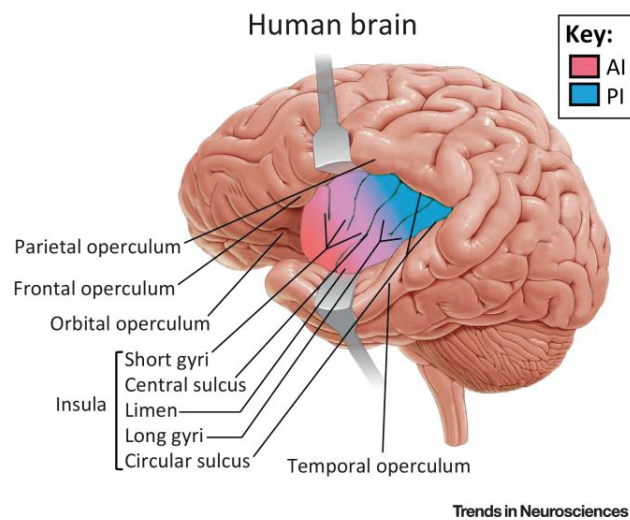


Figure 4. representation of the human insula (Font: Ho Namkung et al., 2017).

In their study Ostrowsky and colleagues (2002) stimulated the insular cortex of patients undergoing a depth stereotactic recording (stereo-electroencephalography, SEEG) during the pre-surgical evaluation of their temporal lobe epilepsies. Their findings showed that both painful and non-painful somesthetic inputs are processed in the posterior part of the insular cortex.

Another clinical work conducted by Krolak-Salmon and colleagues (2003) reported that evoked related potentials (ERPs) to the specific facial expression of disgust were recorded in human insula, specifically in its ventral anterior fields (Krolak-Salmon et al., 2003). The authors used depth electrodes implanted during pre-surgical evaluation of patients with drug-refractory temporal lobe epilepsy and recorded intracerebral event-related potentials to different human facial emotional expressions such as fear, disgust, happiness, surprise, and neutral expression. Their findings highlighted how crucial the ventral anterior insula is in the elaboration of facial emotional expressions, particularly the disgust.

It has been shown that the insular cortex plays a key role in processing visceral sensation, taste, olfactory senses, and in gastrointestinal and respiratory motor functions (Stickler et al., 2003). It has also been argued that this lobe is involved in swallowing and that its electrical stimulation provokes nausea and vomiting in animals (Kaada, 1951). A single case study of a patient presenting with ictal vomiting who underwent bilateral intracranial

exploration including insular depth electrodes was reported by Catenoix and colleagues (2008). Their result showed that the occurrence of ictal vomiting reflect a propagation of the discharge to the insular cortex (Catenoix et al., 2008).

The electrical stimulation of insula has been done also in non-human primates, some of these works are described in the following section.

4.3.2. Electro-stimulation of insula in primates

Several anatomical and microelectrode studies in monkeys showed that primates' insular cortex is involved in processing painful and non-painful somatosensory inputs (Burton and Jones, 1976; Robinson and Burton, 1980; Mufson and Mesulam, 1984; Friedman and Murray, 1986; Apkarian and Hodge, 1989; Hodge and Apkarian, 1990; Schneider et al., 1993; Apkarian and Shi, 1994; Craig et al., 1994; Dostrovsky and Craig, 1996; Blomqvist et al., 2000).

Other studies reported how electrical intracortical microstimulation (ICMS) of single cells in the anterior sector of the insula evokes disgust-related and ingestive behaviors, while stimulation of its dorso-medial sector provokes forelimb movements, and finally stimulation to its mid ventral section produces lip smacking behavior. (Caruana et al., 2011; Jezzini et al., 2012).

Overall, it still remains unclear whether neural activity in the non-human primates anterior insula respond to the vicarious experience and eventually how does it encode empathic experiences.

Another brain region that has been thought to be involved in the elaboration of socially relevant information is the amygdala. In the following section we will give a brief overview of the studies that investigated this limbic brain area by combining electrophysiological recordings with social tasks in monkeys.

4.4. Amygdala: from “body-alarm circuit” to socio-emotional dimension coding and empathy

The amygdala, from the Greek word for almond, has been shown to be a center for fear processing in the brain. This theory was fueled by the numerous experimental works conducted in rodents using fear conditioning paradigm (LeDoux, 2003; Pare et al., 2004). Findings from neurophysiological studies in rodents showed that neurons in the basolateral amygdala (BLA) responded to stimuli associated with fearful events and to the aversive events themselves (LeDoux, 2000; Quirk et al., 1995). (Lau and Salzman, 2009)

However, humans and non-human primates' studies expanded the functional role of the amygdala from primarily vision as “body's alarm circuit” to a center for emotional and social stimuli processing. Indeed, compared to rodents, humans and nonhuman primates' amygdala have an elaborated prefrontal cortex (PFC) many parts of which, especially medial and orbital areas, have extensive and bidirectional connections with the amygdala (Stefanacci et al., 2000, 2002; Ghashghaei et al., 2002, 2007; Wise et al., 2008). The amygdala receives input from a full range of higher sensory and poly-sensory areas to which it projects back in turn, even to primary sensory targets (these connections may be unique to primates) (McDonald et al., 1998; Amaral et al., 2003; Freese et al., 2005). Other brain areas such as the hippocampus, basal ganglia, perirhinal and entorhinal cortices, the basal forebrain, and the hypothalamus are amygdala's output targets (Davis et al., 2000). All these anatomical findings support the idea that the amygdala is implicated in a far more wide-ranging role than “danger alarm”.

Several works showed that amygdala is implicated in reinforcement learning (Holland et al., 1999; LeDoux et al., 2000; Baxter et al., 2002; Everitt et al. 2003; Maren et al., 2004). Lau and Salzman (2009) showed how fugacious epochs of coherent gamma oscillations between amygdala and ventral striatum may be crucial for reinforcement learning (Lau and Salzman, 2009).

To investigate how visual stimuli are linked to values Paton and colleagues (2008) recorded

single neurons in monkeys' amygdala while animals were learning the positive, negative or non-reinforced value of abstract images through a trace conditioning procedure (Paton et al. 2008). After monkeys learned the initial visual stimuli-value associations the value assignments to the images were reversed. Their findings showed that distinct populations of neurons in amygdala encode the positive and negative values of visual stimuli. Moreover, they reported that changes in the values of visual stimuli modulate amygdala's neural activity.

In another study with primates reward-predictive cues were presented to monkeys in different spatial configurations to assess whether amygdala cells encode spatial and motivational information (Peck et al., 2013). The authors reported that amygdala's neural activity was modulated by cue configuration and predicted reward magnitude, and fluctuations in neural activity were correlated with trial to trial variability in spatial attention. The finding that amygdala is a center for the integration of spatial and motivational information suggests that the dysfunction of this limbic area may cause deficits in cognitive processes normally coordinated with emotional responses.

It was previously shown that the amygdala is implicated in the elaboration of various aspects of emotions, memory and social information processing (Cardinal et al., 2002; Phelps, 2006).

Gothard and colleagues (2007) investigated neural activity in the amygdala while monkeys were passively observing images of monkeys' faces, human faces and neutral objects on a computer monitor (Gothard et al., 2007). Human and monkey faces shown to the animals could have threatening or appeasing expressions. Their findings showed that some neurons in the amygdala were responding both to identity and facial expressions, whilst other cells showed pure identity-selective or expression-selective responses. Furthermore they have observed that global activation in the amygdala was larger to threatening faces than to neutral or appeasing faces.

In another study Mosher and colleagues (2014) investigated the neural substrate underlying eye contact in primates. They have recorded the activity of amygdala cells while animals were watching videos of natural behaviors displayed by unfamiliar conspecifics. Their

findings showed the existence of the so called “eye cells”, which are neurons in the amygdala that respond selectively to fixations at the eyes of others and to eye contact. Their findings link the property of the primate amygdala to eye-movements involved in the exploration of the surrounding environments and especially in visual scenes that contain socially and emotionally salient features (Mosher et al., 2014).

Further evidence for the involvement of the amygdala in social processing comes from studies showing that the well known effects of oxytocin (OT) on social approach, trust and attachment in animals (Insel et al., 2001) and humans (Guastella et al., 2008; Kosfeld et al., 2005) could be mediated to a large extent via the connections of the amygdala with OT-rich structures such as the nucleus basalis of Meynert (Freeman et al., 2014; Knobloch et al., 2012).

Although evidences of empathy have been shown from behavioral studies, the neuronal correlates underlying it are still unknown. However, empathic behavior linked to the experience of vicarious reward in primates postulate the existence of mechanisms allowing them to perceive the welfare of others and potentially share their emotional states. This information must be encoded by some brain areas than those described above, and amygdala could be one important candidate.

In a recent study, Chang and colleagues (2015) explored primate amygdala activity while animals were performing a social task in which actor monkeys were presented with three main conditions: reward to self, reward to the partner and reward to a non-live agent (neither) (Chang et al., 2015). These conditions were then combined to make the animals perform choice trials giving them the freedom to cognitively choose the recipient of the reward. Their behavioral results showed how monkeys behaved in a prosocial way by choosing to reward the partner instead of nobody. They recorded cells activity in the basolateral division of the amygdala (BLA) because of its implication in both decision making and social perception. Their behavioral findings confirmed what they have found in their previous study, thus monkeys prosocial behavior. At the neuronal level, a population of BLA cells responded to both obtained and observed rewards, showing similar response scaling as a function of reward size when a monkey chose to grant a reward to itself, to its partner, or to both, but not when the recipient was a non-living agent (neither). Together,

their findings directly implicate the amygdala in social decision making and extend the concept of mirroring to this specific domain. These results stand in contrast with those previously obtained by the same group in the ACCg, where neurons responded in a mirror-like manner but did not show response-scaling to the value of rewards for self and other.

The finding of a brain network that mirrors the vicarious experience of reward brings us very close to the concept of empathy. The theory of action mirroring claims the existence of cells that respond in a similar way to performed and observed actions. These neurons have been mainly described in parietal and motor cortices by the studies conducted by Rizzolatti and colleagues over the past years (Rizzolatti et al., 1987; Rizzolatti and Craighero, 2004; Rizzolatti et al., 2014).

4.5. Mirror Neurons

Mirror neurons (MNs) are a set of visuomotor neurons first discovered in the ventral premotor area F5 of the macaque that discharge both during the execution and observation of goal-directed motor acts (Gallese et al., 1996; Rizzolatti et al., 1996). There is evidence for mirror neurons in two anatomically connected cortical areas in the macaque brain: area F5 in the PMv (ventral premotor cortex) and area PF/PFG in the rostral part of the inferior parietal lobule (Rizzolatti & Craighero, 2004). These cells represent a specific class of motor neurons that discharge both when a monkey performs a motor act and when it observes the same or a similar motor act done by another individual (monkey or human). A fundamental property of mirror neurons is that they respond to the observation of motor acts having the same goal of self-performed motor acts. Mirror neurons discharge during goal-oriented hand actions, such as grasping, tearing and holding. These cells also discharge during ingestive and communicative mouth actions, such as sucking and lip-smacking. The discharge of these cells typically occurs throughout the whole action and is not associated with the contraction of specific muscles. In addition, mirror neurons can fire during actions that are performed with different body parts (Uddin et al., 2007). Mirror neurons also discharge in association with visual and auditory stimuli. A mirror neuron that is active during the execution of a particular action will respond to the sight of similar

actions. For instance, if a mirror neuron discharges during the execution of precision grips, it will also fire when the monkey observes somebody else grasping a small object with a precision grip (Gallese et al., 1996). This pattern of neuronal firing suggests that these neurons code agent-independent actions in rather abstract terms. Functional imaging studies in humans have found a similar circuitry in the human brain (Iacoboni et al., 1999). The human mirror neuron system is typically localized in PMv, dorsal premotor cortex (PMd), and anterior intraparietal sulcus (aIPS; Rizzolatti & Craighero, 2004; Van Overwalle & Baetens, 2009).

There are evidences that monkeys are able to imitate (Ferrari et al., 2006; Subiaul et al., 2004; Voelkl and Huber, 2000, 2007), and it has been suggested that this ability is supported by mirror neurons (Rizzolatti et al., 1999; Rizzolatti and Craighero, 2004; Iacoboni, 2005). Imitation facilitates social interactions, increases connectedness and liking, gets people closer to each other, and fosters mutual care. Following this reasoning, good imitators should also be able in recognizing emotions in other people, which in turn can lead to empathy. Thus, there would be a correlation between the tendency to imitate others and the ability to empathize with them. Several experiments in humans have tested this hypothesis bringing conclusions such as imitation is automatic, linked to liking, and that through imitation and mimicry we are able to feel what other people feel (Chartrand and Bargh, 1999). Other works showed that we are also able to respond compassionately to other people's emotional state (Eisenberg, 2000, , Niedenthal et al. 2005; Tangney et al., 2007; Braten 2007). Based on all these findings, Iacoboni (2009) suggests that one of the main functions of mirror neurons in monkey's brain could be to facilitate others' actions understanding, which could lead to others' actions imitating, and in a final stage to empathize for others.

4.5.1. Mirror system: a circuit for empathy coding?

Empathy is feeling as others feel when observing what they are experiencing. Being empathic is somehow mirroring another person's affective state; thus, in the investigation of a neural system underlying the computation of this feeling, the idea that there might be

‘mirror neurons of empathy’ is a recurrent hypothesis extrapolated from the original descriptions of mirror mechanisms in the premotor and parietal cortices (Gallese and Goldman, 1998; Gallese, 2001; Preston et al., 2002; Decety 2002; Decety and Jackson, 2004; Jabbi et al., 2007; Gazzola et al., 2006; Hojat et al., 2013). Should such neurons exist? In case they do, they should reflect the internal emotional state of others by observing them going through positive or negative experiences. The activity expected from these neurons would be the same or very similar when they respond to experienced and observed experiences. In the next section we will briefly describe what mirror neurons are and how they have been studied till now.

4.5.2. Mirror neurons: sustained and criticized

However, the functional role of mirror neurons has been debated for a long time. Some neuroscientists showed their excitements about this discovery arguing that mirror neurons represent “all that makes us humans!”. In 2000, the neuroscientist Vilayanur Ramachandran made a bold prediction: “mirror neurons will do for psychology what DNA did for biology”. In his book “The Tell-Tale Brain” Ramachandran confirmed his position about the mirror neurons. In the chapter “the neurons that shaped civilization” he gave a large space to mirror neurons arguing that these cells underlie empathy, our capacity to imitate others, accelerate the evolution of the brain, and explain the origin of language. This enthusiasm about mirror neurons was shared by others. The philosopher Grayling sustained that our great gift of empathy is a biologically evolved capacity that lies on the function of mirror neurons.

It has been argued that mirror neurons are crucial in action understanding (Iacoboni, 2009; Gallese et al., 2011). However, findings from medical research conducted in patients with damage to motor networks showed that despite their impairments these individuals were capable to understand others’ actions even though they could not perform them personally. Indeed, “mirror neuron fans” generally accept that action understanding is possible without mirror neurons, but they sustain that these cells allow a deeper understanding of others’ actions. As Iacoboni says in a journal debate (2011), mirror neurons allow “an understanding from within” of others’ actions (Gallese et al., 2011).

This belief was put in discussion by Hickok who clearly sustained that mirror neurons are not the basis for action understanding (Hickok, 2010). He argued that action understanding can be clearly dissociated from “mirror system” and highlighted that the existence of other mechanisms for action understanding is a problem for the mirror neuron theory of action understanding. What Hickok sustained was that the function of mirror neurons is not about understanding others’ actions *per se*, but about using others’ actions to regulate our own actions. Seen this way, mirror neuron activity would be a consequence of action understanding and not viceversa.

Other critics focused on theories claiming that mirror neurons play a central role in human social and cultural evolution by making us empathize with others. Catmur and colleagues (2007), showed that sensorimotor learning experiences can reverse, set aside or magnify mirror-like properties in motor cells, thus mirror neurons are not innate or fixed once acquired. Concerning the role of mirror neurons in cultural evolution the authors sustain that mirror neurons are affected by cultural practices, which can modulate the activity of this category of neurons.

Thus, according to these debates on the functionality of mirror neurons, we aimed to assess whether these cells were the only responsible for our ability to empathize with other individuals.

5. Purpose of the present study

In our work, we investigated how neurons in the anterior insula and amygdala encode self and other's experiences, be them positive or negative. To this purpose, we recorded single cells activity in the AI and amygdala of the monkeys while animals were performing a dynamic social task where they were both involved, in turn, as actors and recipients.

Our behavioral results confirmed what has been found in the work of Ballesta and colleagues (2015). We observed that monkeys were behaving pro-socially with their conspecifics choosing for them mostly positive outcomes instead of aversive ones. In addition to observed evidences of empathic-driven behavior in the animals, we also wanted to investigate the neural substrates underlying such behavior.

We targeted the AI since this brain area has been already shown to be involved in the processing of empathy, especially regarding painful experiences. Indeed, several fMRI works conducted in humans showed that AI responds to both experienced and observed pain (Jackson et al., 2005, 2006; Singer et al., 2006; Lamm et al., 2007; Bird et al., 2010; Hein et al., 2010). All together, these studies suggested that the neural model coding empathy could be the mirror one, accordingly with theories suggesting the mirror model as the ideal one for empathy coding (Iacoboni, 2009).

Other neurophysiological studies in humans and primates targeted the AI. Results obtained by performing electrical intracortical microstimulation (ICMS) of single cells in the anterior sector of the insula in primates evoked disgust related and ingestive behaviors, while stimulation of its dorso-medial sector provoked forelimb movements; finally, stimulation of its mid ventral section produced lip smacking behavior (Caruana et al., 2011; Jezzini et al., 2012). In humans, several clinical works were conducted in epileptic patients using intracranial stimulations (Ostrowsky et al., 2002; Krolak-Salmon et al., 2003; Catenoix et al., 2008). All together, these studies showed insular properties in processing both painful and non-painful somaesthetic inputs, facial emotional expressions, such as disgust, and ictal vomiting.

Thus, so far no work has been done in humans or primates neurophysiology works investigating directly single cells activity in AI in a social context where evidences of empathy towards others were observed. Moreover, fMRI studies, because of the limits of the technique itself, cannot give a neuronal resolution of an investigated brain region activity as electrophysiological recordings can do.

Our aim was to define the nature of “empathy coding neurons” in the AI and to clarify how these cells were coding self and other’s dimension, and the mirroring profile could be one such pattern. However, as explained in the previous sections, the concept of empathy is not only reflecting the other’s emotional state as if it coincides with own internal state, since this kind of processing is the one referred to as pure emotional contagion and compassion. Empathy is also the ability to distinguish self and other’s experiences, as different dimensions of our living reality. Thanks to this distinction, we are able to see things from another’s perspective, an ability that is referred to as “perspective taking theory”, a key component of cognitive empathy (Taylor et al., 1991).

Many of the studies described above showed how monkeys are able to distinguish self from other individuals as agents of performed actions and recipients of possible outcomes (Falcone et al., 2012, 2017; Haroush et al., 2015, Chang et al., 2013, 2015).

We questioned ourselves whether in the AI, a brain region labeled as a neural marker of empathy, there are the neural correlates operating a self-other distinction both in terms of agency: “I am acting” and “you are acting”, and destination: “I am receiving something” and “you are receiving something”. And in case these classes of cells exist, which are their pattern activity?

To assess these questions we designed an imperative task during which both animals were experiencing different type of outcomes for self and other. We inserted the imperative mechanism in this block of the task since it was allowing us to assess neural activities for conditions that monkeys would not have rationally and pro-socially selected in the choice block of the task, such as airpuff to self or other, and avoidance of reward to the partner. Indeed, behavioral results in the choice block revealed that monkeys were mostly choosing to reward the partner instead of sending him the airpuff. Moreover animals were behaving rationally showing that their understanding of the task by selecting for selves the reward instead of the airpuff.

Our neuronal findings report that distinct population of neurons respond specifically to outcomes for self and other, differentiating between their appetitive and aversive nature. Interestingly, the neuronal population responding to the aversive outcome showed mainly three profiles of activity: neuronal representation of conspecifics' unpleasant experience, neuronal representation of own unpleasant experience and a minority of neurons showing mirroring properties between self and other. Our results suppose the existence of two neural substrates in AI processing empathy: a) neurons that selectively distinguish between self and other's unpleasant experience in a consistent way with perspective taking models of empathy; b) neurons that present a different tuning between experienced and observed aversive outcomes, alongside with the minority of the mirror-like cells, which support an emotional resonance model for empathy coding. All together, these results demonstrate how single neurons in AI are involved in the processing of empathy by assessing its multidimensional affective and cognitive nature.

For neural responses to the positive outcome, we found activities exclusively for own small reward and no response to the other's small reward. The 'self selective' cells responded also to self big reward. Notably, some cells responded to other's big reward as well, a salient event in the task, but not to other's small positive outcome. Therefore, the activity for the positive outcomes seems to be modulated by the salience gradient of reward, especially for cells encoding other's outcome.

PART II

EXPERIMENTAL SESSION

1. Materials and Methods

Animals

Three male long-tailed macaques (*Macaca fascicularis*) (monkey 1 aged 8 years, weight 7.5 kg; monkey 2 aged 7.5 years, weight 7 kg; monkey 3 aged 7.5 years, weight 8 kg) were used as subjects. Monkey 1 and 2 were used for both electrophysiological and behavioral recordings while monkey 3 was just used for behavioral recordings. Animals were initially housed as a mini-colony in a large enclosure where they could have direct physical interaction between each other, while during the experiments they were isolated in separated cages. When isolated, monkeys could still communicate vocally and visually thanks to the transparent placards used to separate their individual cages. Behavioral observations of the animals in the cage allowed us to well characterize the hierarchical organization between them: monkey 1 was the dominant while monkey 2 and 3 were the submitted ones. Animals were fed with monkey chow, fresh and dry fruits, vegetables and placed under water restriction with one day of free access to water each week. In order to entertain them when they were isolated and promote social interactions between them during their cohabitation period the cages were enriched with different toys and objects to manipulate.

Animals familiarization with the experimenter and laboratory environment

To lower the level of stress of the animals during the experimental sessions monkeys were trained to interact with the experimenter while they were freely acting in the cages.

A basic type of training named “clicker training” was used to make the monkeys have daily interactions with the experimenter (Gillis et al., 2012). In this training the experimenter was presenting the monkey with an object that the animal was supposed to touch. As soon as the monkey touched the object an auditory stimuli, a *click*, was produced and the animal was rewarded with some juice or a piece of fruit. This kind of training allowed the monkeys to interact with the human subject also outside the laboratory context facilitating the preparative steps before the beginning of each experimental session such as transporting the animals from their cages to the experimental rooms.

Monkeys were taken to the experimental rooms while they were comfortably sitting in primate chairs. Monkeys were guided by the experimenter to sit in their chairs with a metallic bar which was attached to the collars of the animals thanks to a kind of clips at its extremity. Moreover, the passage from their standard cages to the chair was mediated by smaller cages named “transfer cages”. Transfer cages were smaller cages compared to the housing cages and adjacent to these ones. Part of the basic trainings was also to make the animals go spontaneously in the transfer cages without stressing them before bringing them to the labs. The clicker training played an important role in this part as well (once touched the object in the transfer cage animals received generous amount of rewards by going and resting quietly in there).

At the beginning animals were taken in the labs with the experimenter just to get used to the laboratory setting and stayed there in the primate chairs being rewarded occasionally. This helped them to associate the laboratories with the chance for them to obtain rewards and motivated them to go easily in the lab.

Surgery

All experimental procedures were approved by the animal care committee (Department of Veterinary Services, Health & Protection of Animals, permit number 69 029 0401) and the Biology Department of the University Claude Bernard Lyon 1, in conformity with the European Community standards for the care and use of laboratory animals [European Community Council Directive No. 86–609].

The day before the surgery animals were fasting and had unlimited access to water. To prepare the monkeys to the surgery a pre-anesthesia has been done by intramuscular injections of Ketamine (10mg/kg) and of Domitor (25µg/kg). Once the pre-anesthesia was induced monkeys were transported in the surgery room. A catheter was inserted in the saphenous vein to have a direct access for eventual venal injections in case of respiratory or cardiac complications during the surgery. The entire surgery was performed under isoflurane anesthesia (1.0 - 3.0%) for a deep anesthesia of the animals. Monkeys were intubated with an endotracheal tube which was used to help animals breathing and to diffuse isoflurane during the entire surgery.

The entire surgery was performed under sterile conditions. We used betadine for local sterilization of monkey's cranium before the incisions for the craniotomy. All the surgical instruments have been sterilized the day before the surgery with the autoclave machine.

During the entire surgery the heart rate, breathe frequency, blood pressure, oxygen level and body temperature of the animals have been constantly monitored. Animals body temperature was maintained constant thanks to a heated mattress. After the surgery monkeys were de-intubated and brought back to the cages only after clear awakening signs. Once back in the cages they were isolated from the other animals and heated with an infra-red lamp for some hours and even days if necessary.

Head restraint device and recording chamber implantation

We implanted a head-restraint device in titanium to immobilize monkeys head during the neurophysiological recordings and a recording chamber on the left hemisphere (for monkey 1) and on the right hemisphere (for monkey 2) of the animals for our neuronal recordings. The positioning of the recording chamber has been calculated using the stereotaxic coordinates obtained from an IRM of monkeys done before the surgery. Thanks to the head-fixation system we were also able to perform eye tracking of animals to monitor the visual interaction between them.

Behavioral procedures

All experiments have been conducted in a semi-dark room where two head fixed monkeys were seated face to face in a primate chair. Animals were separated from each other and not able to physically interact thanks to the working space placed between them. The working space consisted in a horizontal board (30 cm x 45 cm) on which three aligned buttons were placed. The buttons were spaced with a distance of 2 cm between them. The central button had a different color (yellow) compared to the other two buttons which had the same color (blue). Images were presented above the buttons using a video projector placed above the board. The inclination of the video projector was regulated such as visual cues were projected in a specular way above the three buttons in front of each animal. The entire experimental setup was stabilized by vertical metallic bars on both sides of the working

space board. Behaviorally monkeys were engaged in a social task divided in two separated blocks: an imperative block followed by a choice one.

However, before engaging the monkeys in the final configuration of the imperative and choice blocks of the task we trained them on simpler versions as described in the following section.

Preliminary training

In our final configuration of the task monkeys were facing each other during the entire experimental session. However, the training of the monkeys did not start in presence of each other since the beginning. Indeed initially monkeys were trained alone and separately. We believed that this could help them, in a first moment, to well concentrate in the learning of the different steps of the task.

Button pushing and button holding

The first thing monkeys learned was the simple pushing of the buttons. We presented a central squared cue randomly above all three buttons. Each monkey had its own central squared cue with a specific orientation and color (these cues were then used in the final protocol as a “go signal”). At the beginning, as soon as monkeys were pushing the button corresponding to the squared cue the visual stimuli disappeared and they were rewarded.

After, they learned how to hold the central button once it was pushed down. In this step, monkeys were rewarded only after they hold the central button for a brief interval that we established. We started from 50 ms of holding time and we slowly increased it from session to session till we reached a “basic holding time” of 1000 ms. At an initial step, monkeys were helped to understand the end of the holding time by the disappearance of the visual stimuli after which they could release the button and get the reward. Once monkeys started to master the holding we randomized it by adding, in a not systematic way, 0 ms, 250 ms, 500 ms, 750 ms, 1000 ms, or 1500 ms to the basic holding time of 1000 ms. In this stage, they have to continue to hold the central cue also after the lateral visual stimuli appeared above the corresponding button. We did this in order to avoid making the monkeys act in

an automatic way and also to always have their attention well engaged in performing the task.

Imperative trials

After monkeys learned how to work with the buttons we introduced them to the visual cues associated with the different outcomes.

Firstly, they got used to see the central squared cue that appeared only above the central button. Then, after they hold the central cue for a brief interval (1000 ms), they saw an image appearing above one of the lateral buttons (to the right or to the left of the central button). We had to teach them to select the images in order to understand the outcomes associated with them. They also had to understand that the correct way to select the images, thus to make a “correct trial”, was to wait for the central cue to disappear after the holding time of the central button of 1000 - 2500 ms (as explained above). Once the central cue disappeared, that was the “go signal”, monkeys had a time limit of 1000 ms to move to the left or right lateral button and manually select the image by pushing its corresponding button (the allowed reaction time they had at the beginning was longer than 1000 ms, up to 5000 ms, in order to give them sufficient time to act; we then reduced it gradually based on their behavioral performance). For the selection of the image no holding of the corresponding button was required, but the monkeys had to wait for a delay of 2200 ms before the outcome delivery

Choice trials

Based on their learning skills, after several training sessions with single images, monkeys were introduced to the choice trials. In the choice conditions we paired the same images the monkeys were presented with in the single image trials (imperative trials) and, initially, we gave the animals a generous amount of reaction time (up to 8000 ms) to make their choice. In the final configuration of the protocol, the allowed reaction time in the imperative and choice block was the same (1000 ms), but we found that monkeys needed a larger interval of time to begin the choice trials since there they were cognitively engaged to make choices between two options.

Non-social context

Our task, both in its imperative and choice configuration, was a social task. Monkeys were performing the entire protocol always in presence of their partner and they were performing trials that had specific consequences for themselves only and for their partners only. But the training of the animals did not start in the social context since the beginning.

Monkeys were trained separately in the set up and initially they were presented only with the visual cues associated to the consequences for self: reward self and puff self. Based on their learning skills, after some training sessions, we added the two control conditions to their block of trials: puff to nobody and reward to nobody. We trained them on both imperative and choice conditions when they were working in absence of the partner in the experimental set up.

Social context

Monkeys were introduced to the social context once we observed them performing the task properly in the non-social context.

We started to make them work in the presence of their partner as completely passive at the beginning. They were introduced to the conditions for the other monkeys, other's reward and other's puff, at this stage of the training since they could understand the recipient of these conditions.

Based on their level of performance, we started to make both monkeys work and interact. The understanding of their turn takings improved with the ongoing of the sessions.

Below is described the very final configuration of the task that we used in all our analyzed sessions.

Imperative protocol

In our protocol we used both appetitive and aversive stimuli where appetitive stimuli consisted in finite quantities of juice delivered to the monkeys from a feeder tube positioned close to their mouths and aversive stimuli was a discrete jet of compressed air,

airpuff (4 bars, 60 ms), delivered to the animals from another tube positioned closed to their pupils. We used a gravity-based solenoid device (Crist Instruments) for the control of reward delivery and a tubing system connected to the solenoid device and a pressure gauge to control the airpuff delivery. In the imperative block animals were presented with six different visual cues where two had consequences for the partner sitting in front of them (the social conditions), two had consequences for a non-live agent (the nobody conditions), and two had consequences for themselves (the non-social conditions). Social conditions were: a) rewarding the partner with a drop of juice and b) sending the partner an airpuff; nobody conditions were: a) delivering the reward to an empty container placed in the setup such as both monkeys could clearly see it and b) sending the airpuff in the empty space surrounding the set-up such as it was not tactilely felt from any of the animals present in the room; and finally non-social conditions which were: a) providing themselves some juice and b) sending themselves the airpuff. Each monkey had his own set of images which was different from those presented to his partner both in terms of shape and color. Animals were forced to select each image in order to go on in the block and complete their turn which was marked by a final single trial called “bonanza”. In the bonanza trial, by selecting a bright circle appearing upon the central button, they were rewarded with a big amount of liquid reward (~2000 ms). Each turn was made by 12 trials in which we presented each image on the left or on the right button, plus the ending “bonanza” trial. After one monkey has completed his turn, his partner started to work for his own cycle which followed exactly the same rules. Monkeys alternated each other with the dynamical turn-taking mechanism described for around 10 cycles each (thus in average 120 trials per monkey).

At the beginning of each trial a central cue (a central square having different colors and orientations for each of the monkeys) was presented above the central button which the monkeys had to hold for 1000 - 2500 ms. An image appeared above the left or right button after 1000 ms from the start of the holding of the central button. However, monkeys were forced to hold the central button for the indicated variable time in order to be able to select the presented image afterward. Once the central cue disappeared (‘go signal’), monkeys had an upper reaction time of 1000 ms to correctly perform the trial. If they did not select the target in this interval of time by pushing the corresponding button, the trial was considered as missed and it was presented again to the animal till when it was correctly

selected by him. If the trial was performed correctly the outcome associated with it (reward or airpuff) was delivered after a delay of 2200 ms from the selection of the target.

Choice protocol

The imperative block was followed by a choice one in which we used and paired the same visual cues with which monkeys were presented in the previous block. Monkeys were cognitively engaged in a challenging social decision making task made by four conditions: a) reward the partner by sending him some juice or punish him by sending him the airpuff, b) prevent the partner to get the airpuff by sending it to nobody or send the partner the airpuff, c) offer the partner some juice or send the reward to nobody instead and d) procure himself some reward or send himself an airpuff. Also here animals alternated each other with turn takings and were forced to perform each trial condition in order to go on in the block. Each cycle was made by 8 trials since we presented each couple of images twice by switching the positions of right and left targets compared to the central cue. The end of each cycle was signaled again by the 'bonanza'. Choice trials started with the appearance of a central cue over the central button. Monkeys had to hold the central button for a variable time of 1000 - 2500 ms. After 1000 ms from the beginning of the holding central button two images appeared simultaneously in correspondence of left and right button. Once the central cue disappeared, monkeys were able to make their choice by selecting the image of interest with an allowed reaction time of 1000 ms, otherwise the trial was considered as missed and presented again till it was correctly performed by the animals.

Eye tracker system

Each animal's eye position was monitored using two Eye-Trac 6 (ASL, Bedford MA) infrared video eye trackers (200-Hz sampling rate). At the beginning of each session monkeys' eyes positions were calibrated using a board positioned at 40 cm of distance from monkeys' faces. The surface of the board was marked by a grid of 9 LEDs that helped us to define a precise spatial portion in which monkey's eyes were properly calibrated. Monkeys were trained to fix each of these points by receiving a liquid reward. Once the

calibration was done the board was removed such us the animals could see each other during the performance of the task.

Recording sites definition

At the beginning of each session we defined the precise site of recording in the anterior insula or in the amygdala based on the brain area we were aiming. Both monkeys used for electrophysiological recordings did a functional anatomical IRM after the positioning of their recording chambers. The IRM images were acquired with the Siemens Magnetom Sonata Maestro Class with 1,5 T at CERMEP, a specialized center in Lyon for in vivo image acquisition. Before the scan animals were deeply anesthetized with Ketamine (10mg/kg). Monkeys were head fixed with a stereotaxic system compatible with the MRI machine.

We introduced a guide tube in the center of the grid inserted in the recoding chamber. We used the coordinates of the guide tube to reconstruct the precise areas of anterior insula and amygdala reachable from the holes of the grid positioned inside the recording chamber. A 3D anatomical image was obtained with an acquisition time of 2,89 ms, 1mm slices, 255x255 matrix. Figure 5 shows an example of the MRI obtained from monkey 1. The accessible regions of anterior insula and amygdala are countered in red.

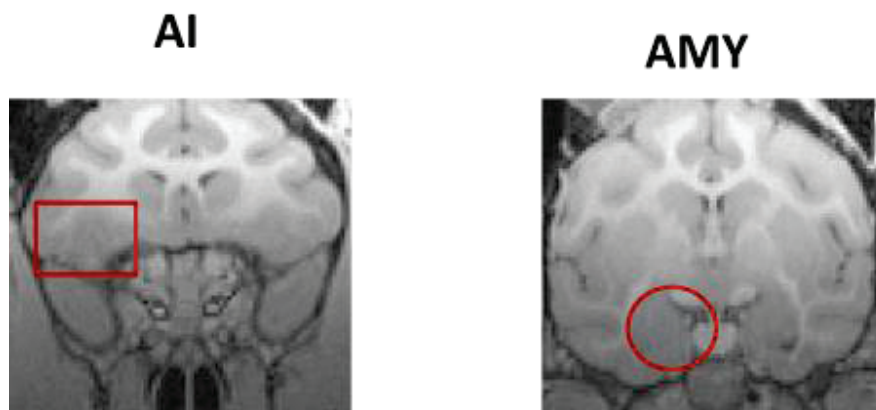


Figure 5. Images from MRI of monkey 1 after the surgery for the positioning of the recording chamber. Anterior Insula (AI) and amygdala (AMY) are counterred in red.

Electrophysiological recordings

Neuronal signals were acquired using 16 channels U-probe (Plexon) tungsten electrodes (15 μ m diameter, impedance 275 \pm 50 k Ω) with an inter-contact spacing of 300 μ m (Figure 6).



Figure 6. 16 contacts U-probe electrode used for neuronal recordings. The channels of the electrode are equally spaced between each other with a distance of 300 μ m, while the distance between the last channel and the point of the electrode is 700 μ m.

Electrodes were connected to a connector and to the Plexon PBX preamplifier (gain of 1000). We used a motorized multi electrode drive (NaN) to bring the probes to the desired depth for our recordings. Spike data were sampled online with a frequency of 20KH and recorded with Spike2 acquisition system. They were then processed offline using Plexon Offline Sorter software to better isolate single cells.

Neuronal and behavioral analyses have been conducted using Matlab R2015 (The MathWorks, Natick, MA). We analyzed only the sessions where neurons were modulated by the task and we focused on the activity of the cells in the imperative block.

2. Scientific contributions

CHAPTER

14

Social Decision-Making in Nonhuman Primates

M. Jazayari^{1,2}, S. Ballesta^{1,2,3}, J.-R. Duhamel^{1,2}

¹Centre National de la Recherche Scientifique, Bron, France; ²Université Lyon 1, Villeurbanne, France; ³The University of Arizona, Tucson, AZ, United States

Abstract

In primates, being able to predict the immediate and long-term consequences of their actions upon others requires keen cognitive abilities that are essential for individual fitness, as inappropriate social behavior can be synonymous with aggression or exclusion from the group. Behavioral and social neurosciences have turned their attention to nonhuman primates to gather new insights into the psychological and biological bases of normal and maladaptive social behavior. In this chapter, we describe experimental studies that have overcome the challenge of using real partners to investigate nonhuman primates' decision-making in a well-controlled social environment. We specifically focus on the neuronal mechanisms of social decision-making and summarize the actual state of our knowledge about the implications of various cortical and subcortical areas in the production of other-regarding behaviors that are naturally involved in the primate's prosocial, cooperative, or competitive interactions.

INTRODUCTION

Primates live in highly social environments characterized by strong competition and dominance hierarchies, but also by cooperation and empathy-driven behaviors that serve to promote social bonds and contribute to group members' fitness. The way we adapt to a dynamic, changing social environment depends on our ability to perceive social stimuli; to convey, but also sometimes conceal, our emotions and intentions; and to keep track of those of others. There has been a growing interest in the identification and characterization of the brain areas dedicated to the processing of social information and in their relation to other large-scale systems such as the cerebral networks involved in reward processing and emotions [1–3].

Compared to other vertebrates, human and nonhuman primates have a larger neocortex and this is

to be directly correlated with the size and complexity of their social systems [4]. Furthermore group size has been shown to drive cortical gray matter expansion in key neocortical areas for social cognition [5]. Obviously, human social behavior has reached an unrivaled degree of sophistication in the course of primate evolution, thanks to some particularly developed cognitive abilities like theory of mind or recursive thinking. This should not lead us to minimize the relevance of comparative studies and of the parallels that can be drawn with nonhuman primate social cognitive abilities. There are pros and cons to the possible homologies a given nonhuman primate species affords with respect to brain architecture, behavioral repertoire, and ecological pressures. Like all models, the macaque monkey has some limitations but it has established itself over the years as a gold standard thanks to the wealth of accumulated knowledge on the neurobiology relevant to social decision-making, including the dopamine reward system, amygdala complex, temporal cortex, insula, and the anterior cingulate (ACC), ventromedial (VM), and orbitofrontal (OFC) subdivisions of the prefrontal cortex (Fig. 14.1). Human patients with VM or OFC lesions have been described as engaging in socially inappropriate behavior [6], lacking awareness of social norm violations [7], failing to adjust their communicative behavior according to the social partner [8], or showing reduced aversion to social inequity [9]. In monkeys, experimental lesions of the OFC similarly point to the role of these structures in a number of social domains, e.g., responding to and producing communicative facial expressions, affiliation and bonding, and aggression [10–12]. In addition, lesions of the ACC can reduce the interest of monkeys in social stimuli [13]. Altered social avoidance, inhibition, and aggression have been found following ablation of the amygdala complex [12,14]. In Chapter 15, Noonan et al. describe the striking

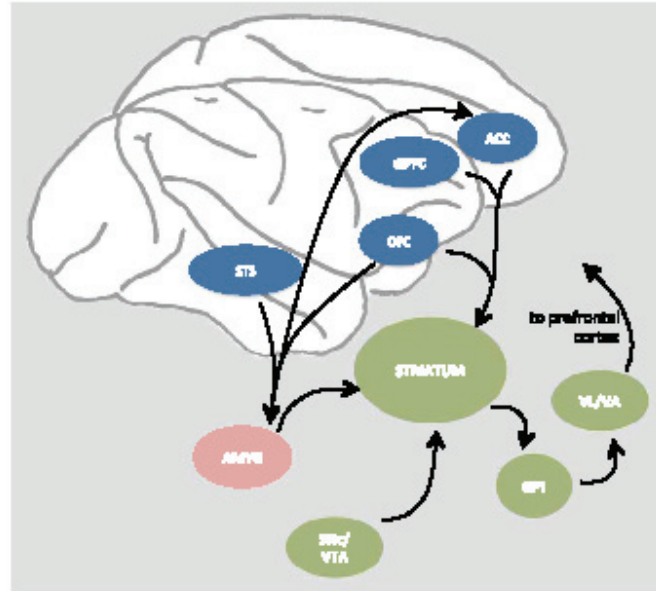


FIGURE 14.1 Simplified schematic representation of the primate reward and social decision-making network. Midbrain nuclei containing dopaminergic neurons: AMYG, amygdala; GPI, internal globus pallidus; SNc/VTA, substantia nigra pars compacta/ventral tegmental area; STS, superior temporal sulcus; VL/VA, ventral lateral and anterior thalamic nuclei. Subdivisions of the prefrontal cortex: ACC, anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; OFC, orbitofrontal cortex.

similarities between the human and the macaque social brains using anatomical and functional brain imaging. In the following sections we draw on studies of macaque behavior and neuronal mechanisms that contribute to our understanding of the social brain. Our focus will be on studies of real dyadic interactions that have examined other-regarding behavior, many of which have relied on social decision-making paradigms as a model for natural prosocial, cooperative, or competitive interactions (see also Chapter 18 for a discussion of strategic reasoning algorithms revealed by simulated social interactions).

BEHAVIORAL STUDIES ON SOCIAL DECISION-MAKING

In their natural habitat, primates face demanding social challenges for which specialized cognitive competencies are assumed to be advantageous, like the capacity to infer what others can hear or see [3,15,16] or to predict others' emotions or intentions [17]. Although higher forms of altruism may require uniquely human mentalizing and moral reasoning

abilities, much evidence supports an evolutionary continuity in the motivational and affective mechanisms that regulate affiliative behaviors [18–21]. Ethological observations show that nonhuman primates are sensitive to others' states, and their ubiquitous social play, grooming behavior, and hormonal correlates suggest an ability to conceive what is pleasant or unpleasant for others [22–24]. Pioneering experimental studies have shown that macaques can perceive and seek to alleviate their peers' distress [25,26], and more recent studies have shown emotional contagion in rodents and its promotion of helping behavior [27–29]. Studies in monkeys have evaluated their capacity to take into account the welfare of a conspecific using simple social dilemmas, such as choosing between one action leading to a reward for self only and another action leading to a reward for self as well as for a passive observer. Despite methodological differences, these experiments consistently show that monkeys often favor the benevolent option, to a degree that depends on factors such as kinship, dominance status, and preexisting social bonds among nonkin ([30–32], see also Chapter 19).

The question arises as to the proximal mechanisms of such prosociality. A strong unifying hypothesis is that

prosocial behavior is shaped by the reinforcing value of certain social events. One proposal is that the mere sight of a conspecific receiving juice might recruit, through some form of mirror mechanism, the same brain reward circuit as an actual drop of juice [31]. Alternatively, it could be that juice allocation increases social attention received from the partner, which in turn initiates a positive loop of rewarding social interaction and promotes prosocial behavior through the monkeys' social attachment system [32,33]. Conversely, observing a conspecific being inflicted an aversive stimulus could be experienced as unpleasant, deactivate reward circuits, and negatively reinforce the antisocial option. The idea that prosocial actions generate their own rewards is related to the warm-glow hypothesis of human altruism: doing good makes us feel good [33,34]. Similarly, empathy

theories postulate that prosocial behavior directed to suffering individuals might primarily aim at alleviating vicariously experienced negative affect [35]. Empirical evidence for an empathy-based account is provided in a study by Ballesta and Duhamel [32], who challenged pairs of cynomolgus and rhesus macaques with social decisions involving appetitive and aversive outcomes while monitoring both animals' oculomotor behavior (Fig. 14.2). The results show that most monkeys take into account the fate of their peer when making choices that could cause outcomes like a drop of juice or an unpleasant air puff being delivered to their partner. Positive correlations between these two types of social decisions indicate that monkeys have coherent motivations in different social interactions. Importantly, benevolence was associated with enhanced social

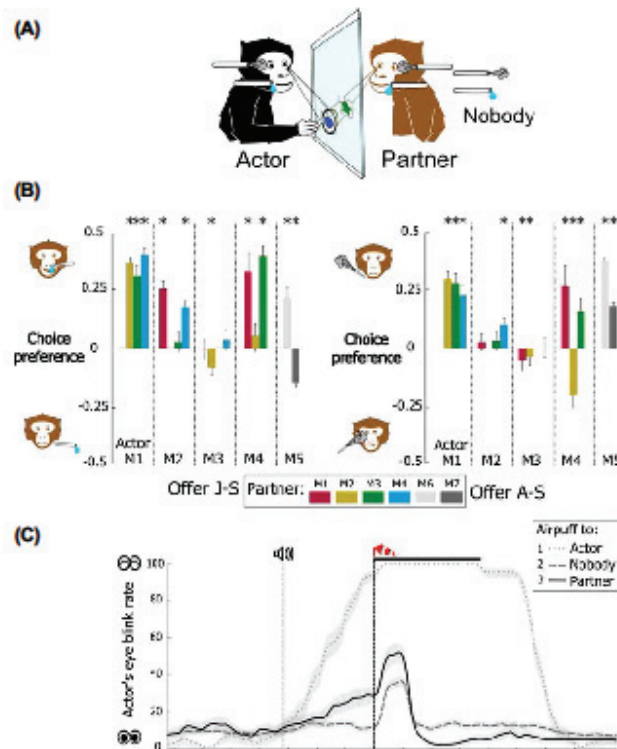


FIGURE 14.2 Evidence of rudimentary empathy during social decisions. (A) Pairs of monkeys sitting face to face made social decisions by touching images on a transparent touch slab. (B) In a majority of dyads, the monkey actor exhibits prosocial behavior, preferring to procure juice rewards and to refrain from administering unpleasant air puffs to the partner. (C) Affective-state matching observed through anticipatory eye blinking on trials in which the actor monkey chose to send an air puff to self or to its partner compared to nobody. Adapted from Ballesta S, Duhamel J-R. Rudimentary empathy in macaques' social decision-making. *Proc Natl Acad Sci USA* 2015;112:15516–21.

III. SOCIAL DECISION NEUROSCIENCE

communication through deliberate mutual gaze and with empathic eye blinking in anticipation of the partner's being inflicted an aversive stimulus (Fig. 14.2C). Such enhanced eye blinking when watching a peer receiving an air puff is highly suggestive of induced negative affect in the observer, and in fact indifference or malevolence was associated with lower or suppressed such responses.

Theories of empathy, broadly defined as vicarious experiences of the affective states of others [19,21], postulate that shared representations serve as a proximate mechanism sustaining altruistic behaviors [18–20,36]. Our finding of similar behavioral responses to the direct experience of an aversive stimulus and to the observation of its impact upon a peer would thus appear consistent with simulation theories of empathy. Finally, it is noteworthy that the small population of tested macaques does not behave in a uniformly benevolent manner, but showed consistent partner preferences. Using manual and automatic scoring of spontaneous interactions in their living space [37], we assessed the social organization and affiliative behavior patterns of monkeys and found that benevolent ones spend more time and exchange more grooming with their partner than nonbenevolent monkeys. This is in accordance with ethological observations of social behaviors such as coalition building or reconciliations [38,39], which emphasize that nonhuman primates actively establish and preserve preferential social relationships. Together, these results show that macaques' responses to the welfare of others can be measured objectively, opening the way to the investigation of the neuronal processing during social interactions in a laboratory setting.

NEURONAL CORRELATES OF DECISION-MAKING IN A SOCIAL CONTEXT

The neural substrates underlying social behavior and motivation systems are likely to intersect for several obvious reasons. A first reason is that social information possesses intrinsic rewarding qualities. The tremendous success of smartphone photo exchange apps and online social networks is a patent testimony to the high value we assign to observing and learning about our conspecifics. This is true also of monkeys who spontaneously gaze at socially significant features of visual scenes [40–42] and are willing to partially sacrifice a primary reward such as fruit juice to be allowed to view images of conspecifics [43]. This has led to the view that, in addition to temporal lobe systems specialized for the processing of social information such as faces, and particularly the gaze direction, emotional expressions, and identity of faces [44–47], other regions participate in encoding the subjective value of such social stimuli [43].

Another obvious reason, and this is what the following section will focus on more specifically, is that vital resources come in limited amounts and thus generate different social adaptations to deal with competition and dominance status of peers. Access to goods (food) and services (grooming by a partner) in a social context implies that "others" can be used as comparison points and directly affect perception of one's own welfare.

Numerous electrophysiological experiments have investigated the neuronal correlates of motivated behaviors and reward processing in monkeys, notably in the dopamine reward system (see Ref. [48] for a review). The representation of various natural reinforcers, both appetitive [49,50] and aversive [51], has also been described in the OFC, and activity in this region is described as reflecting the subjective value of rewards during decision-making [52,53]. To determine whether neural encoding of motivational value reflects only the relation between a physiological state (e.g., blood glucose concentration) and a physical magnitude (number of calories supplied by a food item), or whether it encompasses a broader, more psychological dimension of value, the effect of social context on reward-related activity has been investigated in the OFC and in related structures.

In an attempt to extend and to investigate the role of the OFC from value-based decisions to socially motivated ones, Azzi et al. [54] trained monkeys to produce a manual response to visual cues predicting subsequent juice rewards for themselves and for passive observers (Fig. 14.3). The results showed that the spontaneous spike discharge activity of the OFC neuronal population is globally enhanced in a social context compared to a nonsocial one. Furthermore, value-sensitive neurons were found to respond to a fixed-size reward for self with different discharge rates depending on whether monkeys were working to obtain a reward for self only or for both self and a partner monkey. This is illustrated for a single unit and for a population of reward-value coding cells responding with a monotonic increase in discharge rate as a function of reward size during a series of nonsocial trials (left panels of Fig. 14.3B and C). In a social trial block, neuronal firing to a predicted reward for self was modified on trials in which an identical reward was provided to one of the monkey's partners, P1 or P2. Interestingly, these neuronal responses reflected the subjective value of rewards, as inferred from the behavior of the monkeys who, in this particular context, showed less motivation (i.e., performed fewer correct trials) on "shared" than on "unshared" reward trials. Finally, a session-by-session analysis further showed that OFC activity tracked the monkey's current social preferences, expressed in the relative willingness to procure rewards to one or the other partner. Together,

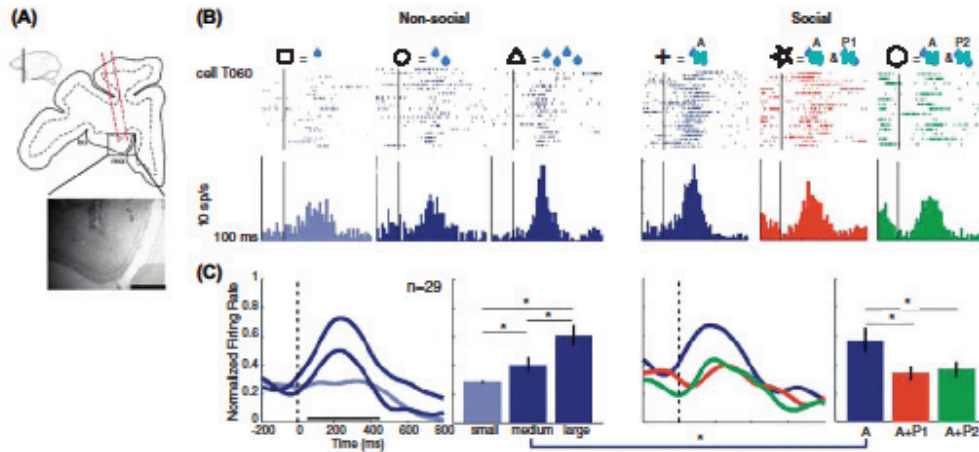


FIGURE 14.3 Social modulation of reward-value signals in the macaque OFC. (A) Single units were recorded in area 13 while monkeys worked to obtain rewards for self only or self plus a passive partner. (B) A positively coding value-sensitive single-unit responding as a function of reward size in a nonsocial context (left rasters and histograms) shows a reduced response in a social context when either partner P1 or partner P2 received a concomitant reward (right rasters and histograms). (C) Population activity of value-sensitive orbitofrontal cortex neurons. *Adapted from Azzi JCB, Sirigu A, Duhamel J-R. Modulation of value representation by social context in the primate orbitofrontal cortex. Proc Natl Acad Sci USA 2012;109:2126–31.*

these results indicate that social influence of reward sharing and partner identity is taken into account in the valuation process operated by the OFC.

Such properties suggest that the OFC could be involved in the neural computations taking place during social comparisons. However, the study by Azzi et al. did not address directly the issue of coding of subjective equity and inequity between one's own welfare and that of other individuals. People naturally compare their resources with those of others and, as Karl Marx noted, perceived reward differences are a major driving force in the pursuit of wealth and a source of perpetual frustration: "A house may be large or small; as long as the neighboring houses are likewise small, it satisfies all social requirement for a residence. But let there arise next to the little house a palace, and the little house shrinks to a hut. The little house now makes it clear that its inmate has no social position at all to maintain, or but a very insignificant one; and however high it may shoot up in the course of civilization, if the neighboring palace rises in equal or even in greater measure, the occupant of the relatively little house will always find himself more uncomfortable, more dissatisfied, more cramped within his four walls" [55]. Sensitivity to unequal distribution of rewards has been well documented in human adults [56] and children [57]. And in a similar and not surprising way it has been shown that also various nonhuman

primate species respond adversely to social inequity [58–62]. Báez-Mendoza et al. [63] investigated neuronal coding of inequity in the ventral striatum using an imperative reward-giving task in which monkeys experienced mainly four different reward conditions: reward to self, to the partner, to both animals, or to nobody. Behaviorally, monkeys exhibited unfavorable inequity aversion. At the single-cell level, as expected from prior studies [64–66], a sizeable proportion of striatal neurons encoded reward to self. However, a subset of striatal neurons were sensitive to reward inequity, coding either disadvantageous inequity (receiving less reward than the partner) or advantageous inequity (receiving more reward than the partner), and more rarely both forms of inequity.

So far we have described brain signals that track changes in subjective utility of own rewards as a function of the social context. Neurons in the striatum can also respond to another individual's actions, but only when resulting in own reward [67]; hence their "frame of reference," like in the OFC, remains self-centered and does not describe directly another individual's reward experience. Vicarious reward- and empathy-based accounts of prosocial behavior postulate the existence of mechanisms allowing an individual to perceive the welfare of others and potentially assign it a value based on self-experience. Such information could be encoded by other brain structures.

III. SOCIAL DECISION NEUROSCIENCE

The amygdala is one such candidate, as it is implicated in various aspects of emotions, memory, and social information processing [1,68]. Primate amygdala neurons respond when animals observe emotional facial expressions of other monkeys [45] and engage in eye contact [45,69]. Furthermore the well-known effects of oxytocin (OT) on social approach, trust, and attachment in animals [70] and humans [71,72] could be mediated to a large extent via the connections of the amygdala with OT-rich structures such as the nucleus basalis of Meynert [73,74]. Chang et al. [75] reported that value-sensitive neurons in the macaque basolateral amygdala do not distinguish between obtained or observed rewards, showing similar response scaling as a function of reward size when a monkey chooses to grant juice to itself, its partner, or to itself and its partner, but not when the recipient was an inanimate agent (a drop of juice falling into a container). The authors thus argue for a motivational mirroring mechanism allowing the evaluation of rewards for self and others in a "common currency."

Social information about agency and reward destination also appears to be processed by the amygdala, as shown by preliminary findings from our laboratory. This region is known to be involved in the learning of associations between visual stimuli and reward or punishment [76]. Fig. 14.4 illustrates a response pattern that

we recorded in the amygdala while two monkeys alternated in performing a reward allocation task. On some trials, monkeys responded to a visual cue associated with juice rewards to self, the partner, or nobody by a manual key press. This example neuron showed a brisk, short latency visual response to cues predicting rewards for the monkey's partner and for nobody (the same empty container condition as in the Chang et al. study), but it was inhibited by a cue predicting a reward for self. Interestingly, it responded much more strongly when the monkey was in the observer's than in the actor's role, hence providing information about the identity of the donor. Such signals could be interpreted as tracking the agent empowered with controlling access to the "resources," which in a certain sense can be considered as a proxy for social dominance [77]. Interestingly, the amygdala has been found to be part of a restricted set of brain areas showing a correlation between gray matter volume and position in the social hierarchy [78]. Other studies have shown that activity in the ACC distinguishes between self and other as the reward recipient [79], and medial prefrontal neurons were found to track rewards and errors made by monkey partners [80]. Taken together, these data indicate that the amygdala and the medial prefrontal cortex carry signals relevant for evaluating other's welfare, action, and social status.

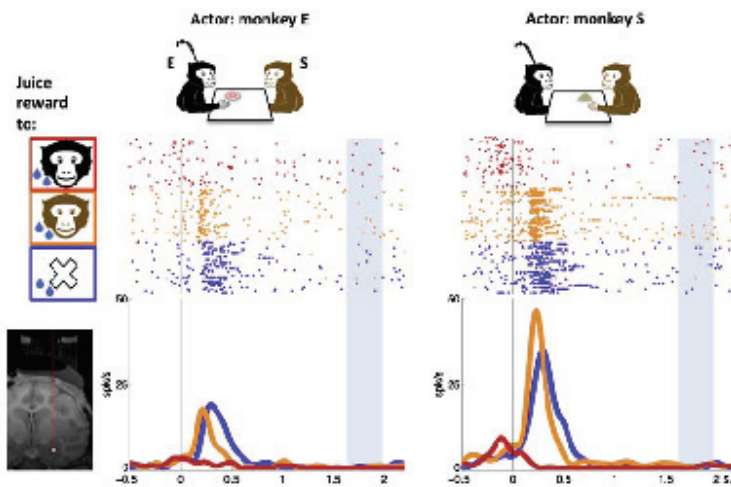


FIGURE 14.4 Reward recipient and agency coding in the macaque amygdala. Two monkeys alternated in performing an imperative task by touching a cue image predicting whether a juice reward would be granted to self, the partner, or nobody. Individual trial rasters (top) and mean spike density (bottom) for a single neuron recorded in the brain of monkey E, while either it or the partner monkey S performed the task, are shown. The neuron responds to cues predicting a reward to monkey S or to nobody, but not to monkey E, and responds more strongly when monkey E observed the task being performed rather than when performing it. Neuronal activity is aligned on cue onset, blue-shaded rectangles represent approximate time of outcome delivery.

These brain structures can thus provide key information about others that is required for empathy-based, prosocial behavior, but also, in other situations, for competition and cooperation. Haroush and Williams [81] used an iterated prisoner's dilemma game to investigate the capacity of monkeys to coordinate their actions. This game mimics a type of social interaction in which individuals must try to anticipate each other's move to pick the most beneficial, or least costly, course of action. When played multiple times, players can respond to each other's actions, retaliate against or forgive noncooperative behavior, and ideally, but not necessarily, converge on the mutually advantageous dominant strategy. The authors found that neuronal activity in the dorsal ACC accurately predicted other player's intention to cooperate or defect, and disruption of ACC activity by electrical stimulation reduced the preference for choosing the mutually beneficial option. In a context of virtual competition, in which monkeys played a video shooting game against each other, Hosokawa and Watanabe [82] identified neurons in the lateral prefrontal cortex (LPFC) that discriminated between task situations, responding more strongly to reward outcomes when playing against a monkey opponent than a computer opponent or no opponent at all. In a follow-up study, the same authors examined the neural mechanisms of competitive behavior under various reward contingencies. An interesting contrast was made between the usual outcome of a competition, in which the winner obtains the reward and the loser obtains nothing, and an egalitarian outcome in which both players are rewarded. Monkeys are obviously not natural-born collectivists, as their performance declined strikingly under the egalitarian reward condition [83]. At the neuronal level, the LPFC discriminated between the two competitive situations and showed little reward-related activity associated with winning in the egalitarian context. Together, these results suggest an important role for the ACC and LPFC in decision-making during competitive social interactions.

CONCLUSIONS AND PERSPECTIVES

The use of laboratory tasks mimicking real-world multiagent interactions of nonhuman primates is a relatively new direction in behavioral neuroscience. Although much work still has to be done, this approach highlights the role of several key players in social motivation and reward evaluation, and provides a fine-grained analysis of the neuronal computations performed by these structures. The picture remains patchy, in part because different experimental tasks have been used to study different areas. The specificity and the information flow among these areas are also far from

being elucidated. Nevertheless a rough sketch of the functional properties described in this chapter can be proposed. The OFC and striatum generate somewhat similar value signals that are weighted by social context, partner preference, and resource distribution and that may provide part of the motivational basis for social tolerance, approach, and avoidance. Value signals are also present in the amygdala, in which a class of cells responding to the value of both obtained and given rewards might be involved in affective state matching, social reward, and prosocial behavior. Finally, neurons in prefrontal areas like the ACC and LPFC, but also in the amygdala, distinguish between attributes of self and other, and represent information about agency and others' current intentions during social decision-making and could thus mediate complex decisions, such as whether and when to engage in cooperative or competitive behavior. The cognitive challenges that are typically being used in these experiments are simplified versions of the choices monkeys have to make when interacting with their conspecifics, and provide even more simplified models of complex human social cognitive processes. However, important features of primates' social decision-making are emerging, such as the fact that it does not necessarily follow a strict utility hypothesis and that everyday social decisions seem to rely, more than any other type of decisions, on affective processing. Indeed, it seems that nonhuman primates are able to create an elaborated representation of their social environment and thus express social preferences that reflect their desire to strengthen or weaken each unique social bond. Primates thus have expectations about the consequences of their behavior upon others' affect, but it is unclear to what extent nonhuman primates "know" about mental states of other live agents in the sense that humans do. Nevertheless, if one admits that a cognitive skill like theory of mind is not simply something one does or does not possess, but is assembled from more elementary abilities like social awareness and emotional empathy, monkeys clearly display some of these competences, at least in rudimentary form. Pursuing the study of the distal and ultimate bases of nonhuman primate's social decision-making will surely participate in fostering a unifying theory dealing with the roots of maladaptive and normal social behavior of humans.

References

- [1] Cardinal RN, Parkinson JA, Hall J, Everitt BJ. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* 2002;26:321–52.
- [2] Haber SN, Knutson B. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 2010;35:4–26.

- [3] Rushworth MF, Mars RB, Sallet J. Are there specialized circuits for social cognition and are they unique to humans? *Curr Opin Neurobiol* 2013;23.
- [4] Dunbar RIM, Shultz S. Evolution in the social brain. *Science* 2007; 317:1344–7.
- [5] Sallet J, Mars RB, Noonan MP, Andersson JL, O'Reilly JX, Jbabdi S, et al. Social network size affects neural circuits in macaques. *Science* 2011;334:697–700.
- [6] Bechara A, Damasio AR, Damasio H, Anderson SW. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 1994;50:7–15.
- [7] Beer JS, John OP, Scabini D, Knight RT. Orbitofrontal cortex and social behavior: integrating self-monitoring and emotion-cognition interactions. *J Cogn Neurosci* 2006;18:871–9.
- [8] Stolk A, D'Imperio D, di Pellegrino G, Toni I. Altered communicative decisions following ventromedial prefrontal lesions. *Curr Biol* 2015;25:1469–74.
- [9] Moretti L, Dragone D, di Pellegrino G. Reward and social valuation deficits following ventromedial prefrontal damage. *J Cogn Neurosci* 2009;21:128–40.
- [10] Babineau BA, Bliss-Moreau E, Machado CJ, Toscano JE, Mason WA, Amaral DG. Context-specific social behavior is altered by orbitofrontal cortex lesions in adult rhesus macaques. *Neuroscience* 2011;179:80–93.
- [11] Machado CJ, Bachevalier J. Measuring reward assessment in a semi-naturalistic context: the effects of selective amygdala, orbital frontal or hippocampal lesions. *Neuroscience* 2007;148:599–611.
- [12] Machado CJ, Bachevalier J. The impact of selective amygdala, orbital frontal cortex, or hippocampal formation lesions on established social relationships in rhesus monkeys (*Macaca mulatta*). *Behav Neurosci* 2006;120:761.
- [13] Rudebeck PH, Buckley MJ, Walton ME, Rushworth MFS. A role for the macaque anterior cingulate gyrus in social valuation. *Science* 2006;313:1310–2.
- [14] Amaral DG, Bauman MD, Capitanio JP, Lavenex P, Mason WA, Mauldin-Jourdain ML, et al. The amygdala: is it an essential component of the neural network for social cognition? *Neuropsychologia* 2003;41:517–22.
- [15] Flombaum JL, Santos LR. Rhesus monkeys attribute perceptions to others. *Curr Biol* 2005;15:447–52.
- [16] Santos LR, Nissen AG, Ferrugia JA. Rhesus monkeys, *Macaca mulatta*, know what others can and cannot hear. *Anim Behav* 2006;71:1175–81.
- [17] Cheney D, Seyfarth R, Smuts B. Social relationships and social cognition in nonhuman primates. *Science* 1986;234:1361–6.
- [18] Decety J. The neuroevolution of empathy: neuroevolution of empathy and concern. *Ann NY Acad Sci* 2011;1231:35–45.
- [19] Preston SD, de Waal FB. Empathy: its ultimate and proximate bases. *Behav Brain Sci* 2002;25:1–20.
- [20] de Waal FBM. The antiquity of empathy. *Science* 2012;336:874–6.
- [21] Batson CD. These things called empathy: eight related but distinct phenomena. In: Decety J, Ickes W, editors. *Soc Neurosci. Empathy*. Cambridge (MA): MIT Press; 2009. p. 3–15.
- [22] Keverne EB, Martensz ND, Tuite B. Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology* 1989;14:155–61.
- [23] Shimada M. Social object play among young Japanese macaques (*Macaca fuscata*) in Arashiyama, Japan. *Primates* 2006;47:342–9.
- [24] Shutt K, MacLarnon A, Heistermann M, Semple S. Grooming in Barbary macaques: better to give than to receive? *Biol Lett* 2007; 3:231–3.
- [25] Masserman JH, Wechkin S, Terris W. "Altruistic" behavior in rhesus monkeys. *Am J Psychiatry* 1964;121:584–5.
- [26] Miller RE, Banks Jr JH, Kuwahara H. The communication of affects in monkeys: cooperative reward conditioning. *J Genet Psychol* 1966;108:121–34.
- [27] Bartal IB-A, Decety J, Mason P. Empathy and pro-social behavior in rats. *Science* 2011;334:427–30.
- [28] Atsak P, Orre M, Bakker P, Cerliani L, Rooszendaal B, Gazzola V, et al. Experience modulates vicarious freezing in rats: a model for empathy. *PLoS One* 2011;6:e21855.
- [29] Ben-Ami Bartal I, Rodgers DA, Bernardor Sarria MS, Decety J, Mason P. Pro-social behavior in rats is modulated by social experience. *Elife* 2014;3:e01385.
- [30] Massen JMM, van den Berg LM, Spruijt BM, Sterck EHM. Generous leaders and selfish underdogs: pro-sociality in despotic macaques. *PLoS One* 2010;5:e9734.
- [31] Chang SWC, Winecuff AA, Platt ML. Vicarious reinforcement in rhesus macaques (*Macaca mulatta*). *Front Neurosci* 2011;5.
- [32] Ballesta S, Duhamel J-R. Rudimentary empathy in macaques' social decision-making. *Proc Natl Acad Sci USA* 2015;112:15516–21.
- [33] Preston SD. The origins of altruism in offspring care. *Psychol Bull* 2013;139:1305–41.
- [34] Aknin LB, Hamlin JK, Dunn BW. Giving leads to happiness in young children. *PLoS One* 2012;7:e39211.
- [35] Jackson PL, Meltzoff AN, Decety J. How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage* 2005;24:771–9.
- [36] Gallese V. Before and below "theory of mind": embodied simulation and the neural correlates of social cognition. *Philos Trans R Soc B Biol Sci* 2007;362:659–69.
- [37] Ballesta S, Raymond G, Pozzobon M, Duhamel J-R. A real-time 3D video tracking system for monitoring primate groups. *J Neurosci Methods* 2014;234.
- [38] Cordis M. Post-conflict reunions and reconciliation in long-tailed macaques. *Anim Behav* 1992;44:57–61.
- [39] Widdig A, Streich WJ, Tembrock G. Coalition formation among male Barbary macaques (*Macaca sylvanus*). *Am J Primatol* 2000; 50:37–51.
- [40] Nahm FK, Perret A, Amaral DG, Albricht TD. How do monkeys look at faces? *J Cogn Neurosci* 1997;9:611–23.
- [41] Dahl CD, Wallraven C, Bulthoff HH, Logothetis NK. Humans and macaques employ similar face-processing strategies. *Curr Biol* 2009;19:509–13.
- [42] Mosher CP, Zimmerman PE, Gothard KM. Videos of conspecifics elicit interactive looking patterns and facial expressions in monkeys. *Behav Neurosci* 2011;125:639–52.
- [43] Watson KK, Platt ML. Social signals in primate orbitofrontal cortex. *Curr Biol* 2012;22:2268–73.
- [44] Perrett DI, Hietanen JK, Oram MW, Benson PJ. Organization and functions of cells responsive to faces in the temporal cortex. *Philos Trans R Soc Lond B Biol Sci* 1992;335:23–30.
- [45] Gothard KM, Battaglia FP, Erickson CA, Spitzer KM, Amaral DG. Neural responses to facial expression and face identity in the monkey amygdala. *J Neurophysiol* 2007;97:1671–83.
- [46] Sliwa J, Duhamel J-R, Pascalis O, Wirth S. Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proc Natl Acad Sci USA* 2011;108:1735–40.
- [47] Sliwa J, Planté A, Duhamel J-R, Wirth S. Independent neuronal representation of facial and vocal identity in the monkey hippocampus and inferotemporal cortex. *Cereb Cortex* 2016;26:950–66.
- [48] Schultz W. Neuronal reward and decision signals: from theories to data. *Physiol Rev* 2015;95:853–951.
- [49] Rolls ET, Hornak J, Wade D, McGrath J. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J Neurol Neurosurg Psychiatry* 1994;57: 1518–24.
- [50] Rolls ET. The orbitofrontal cortex and reward. *Cereb Cortex* 2000; 10:284–94.
- [51] Morrison SE, Salzman CD. The convergence of information about rewarding and aversive stimuli in single neurons. *J Neurosci* 2009; 29:11471–83.

III. SOCIAL DECISION NEUROSCIENCE

- [52] Tremblay I, Schultz W. Relative reward preference in primate orbitofrontal cortex. *Nature* 1999;398:704–8.
- [53] Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature* 2006;441:223–6.
- [54] Azzi JCB, Sirigu A, Duhamel J-R. Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc Natl Acad Sci USA* 2012;109:2126–31.
- [55] Marx K. Wage labor and capital. *Neue Rheinische Zeitung*; 1849.
- [56] Adams JS. Inequity in social exchange. *Adv Exp Soc Psychol* 1965;2.
- [57] LoBue V, Nishida T, Chiong C, DeLoache JS, Haidt J. When getting something good is bad: even three-year-olds react to inequality. *Soc Dev* 2011;20:154–70.
- [58] Brosnan SE. Justice and fairness-related behaviors in nonhuman primates. *Proc Natl Acad Sci USA* 2013;110:10416–23.
- [59] Brosnan SE, de Waal FBM. Monkeys reject unequal pay. *Nature* 2003;425:297–9.
- [60] Burkart JM, Fehr E, Efferson C, Van Schaik CP. Other-regarding preferences in a non-human primate: common marmosets provision food altruistically. *Proc Natl Acad Sci USA* 2007;104:19762.
- [61] Massen JMM, Van Den Berg LM, Spruijt BM, Sterck EHM. Inequity aversion in relation to effort and relationship quality in long-tailed Macaques (*Macaca fascicularis*). *Am J Primatol* 2012;74:145–56.
- [62] Proctor D, Williamson RA, de Waal FBM, Brosnan SE. Chimpanzees play the ultimatum game. *Proc Natl Acad Sci USA* 2013;110:2070–5.
- [63] Biazé-Mendoza R, van Coeverden CR, Schultz W. A neuronal reward inequity signal in primate striatum. *J Neurophysiol* 2016;115:68–79.
- [64] Hikosaka O, Sakamoto M, Usui S. Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J Neurophysiol* 1989;61:814–32.
- [65] Apicella P, Sarnati E, Ljungberg T, Schultz W. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J Neurophysiol* 1992;68:945–60.
- [66] Cromwell HC, Schultz W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* 2003;89:2823–38.
- [67] Biazé-Mendoza R, Schultz W. The role of the striatum in social behavior. *Front Neurosci* 2013;7.
- [68] Phelps EA. Emotion and cognition: insights from studies of the human amygdala. *Annu Rev Psychol* 2006;57:27–53.
- [69] Mosher CP, Zimmerman PE, Gofhard KM. Neurons in the monkey amygdala detect eye contact during naturalistic social interactions. *Curr Biol* 2014;24:2439–64.
- [70] Insel TR, Young LJ. The neurobiology of attachment. *Nat Rev Neurosci* 2001;2:129–36.
- [71] Guastella AJ, Mitchell PB, Mathews E. Oxytocin enhances the encoding of positive social memories in humans. *Biol Psychiatry* 2008;64:256–8.
- [72] Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E. Oxytocin increases trust in humans. *Nature* 2005;435:673–6.
- [73] Freeman SM, Inoue K, Smith AL, Goodman MM, Young LJ. The neuroanatomical distribution of oxytocin receptor binding and mRNA in the male rhesus macaque (*Macaca mulatta*). *Psychoneuroendocrinology* 2014;45:128–41.
- [74] Knobloch HS, Charlet A, Hoffmann LC, Eliava M, Khrulev S, Cetin AH, et al. Evoked axonal oxytocin release in the central amygdala attenuates fear response. *Neuron* 2012;73:553–66.
- [75] Chang SWC, Fagan NA, Toda K, Utevsky AV, Pearson JM, Platt ML. Neural mechanisms of social decision-making in the primate amygdala. *Proc Natl Acad Sci USA* 2015;112:16012–7.
- [76] Sanghera MK, Rolls ET, Roper-Hall A. Visual responses of neurons in the dorsolateral amygdala of the alert monkey. *Exp Neurol* 1979;63:610–26.
- [77] Prutau C, Voelkl B, Van Damme E, Noë R. Supply and demand determine the market value of food providers in wild vervet monkeys. *Proc Natl Acad Sci USA* 2009;106:12007–12.
- [78] Noonan MP, Sallet J, Mars RB, Neubert FX, O'Reilly JK, Andersson JL, et al. A neural circuit covarying with social hierarchy in macaques. *PLoS Biol* 2014;12:e1001940.
- [79] Chang SWC, Gariépy J-F, Platt ML. Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci* 2012;16:243–50.
- [80] Yoshida K, Saito N, Iriki A, Isoda M. Social error monitoring in macaque frontal cortex. *Nat Neurosci* 2012;15:1307–12.
- [81] Haroush K, Williams ZM. Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* 2015;160:123–45.
- [82] Hosokawa T, Watanabe M. Prefrontal neurons represent winning and losing during competitive video shooting games between monkeys. *J Neurosci* 2012;32:7662–71.
- [83] Hosokawa T, Watanabe M. Egalitarian reward contingency in competitive games and primate prefrontal neuronal activity. *Front Neurosci* 2015;9:165.

Title: Neuronal basis of empathy in the primate insula.

Authors: Mina Jazayeri, Sophie Gilardeau, Rossella Cirillo, Sylvia Wirth, Jean-René Duhamel*

Affiliations: Institut des Sciences Cognitives Marc Jeannerod, Université de Lyon 1 and Centre National de la Recherche Scientifique

*Correspondence: duhamel@isc.cnrs.fr

This word file includes:

Abstract

Main Text

References

Acknowledgements

Figures

Figures Legends

Successful Submission of a Manuscript to Science

Manuscript Title: Neuronal basis of empathy in the primate insula

Author: Jazayeri

Manuscript Number: aar2263

Abstract

Adaptive social interaction is based on our ability to understand the intentional and affective states of others. Brain imaging studies in humans claimed that anterior insula (AI) encodes empathy through its activation to both observed and experienced pain. Evidences of empathy in monkeys have been reported, but the neural bases of this behavior have yet to be assessed. We recorded single-unit activity in the macaque AI while animals performed a social task where pleasant or unpleasant stimuli could be directly experienced or observed. Our results showed that neurons responded distinctly to self and other's aversive experience, and rarely to both. The majority of neurons encoding positive stimuli responded exclusively to own outcome. However, some cells showing mirror-like properties responded to both self and other's highly salient positive outcomes. We suggest a neuronal model of empathy that accounts for the distinctive features between "feeling" and "feeling for".

Empathy is a multifaceted emotion that can be loosely defined as an affective state triggered by a similar state in another individual. It is sometimes accompanied by physiological and motor responses such as stress contagion and motor mimicry and is considered to facilitate prosocial behaviors. The ability to understand other's feelings has been assumed to involve shared mental representations between own and other's affective states (1, 2). Other theories emphasize cognitive features, like perspective-taking and mentalizing processes that for some authors might be uniquely human (3). Yet a clear case for evolutionary continuity is present in many species that display emotional contagion, empathic concern, and engage in consolation and helping behavior (4–8), thus offering an opportunity to investigate the basic neural mechanisms of social emotions.

The anterior insula (AI) is considered to play a central role in empathy. It is involved in processing afferent bodily signals and has strong connections to limbic areas (9). In humans, it is activated by felt and observed pain (10), suggesting that the AI represents a common neural substrate for the shared subjective emotional features between nociceptive and empathic pain. A question left unanswered by neuroimaging studies is whether the overlap between real and vicarious experiences observed at the scale of a cortical area reflects a similar overlap at the single neuron level. Mirror neurons discovered in macaques' brain discharge during both execution and observation of specific motor behaviors (11). The function of mirror neurons is still debated. Initially proposed as a mechanism for other's action recognition, imitation and prediction during social interaction (12), different authors have extended its function to empathy and understanding of other's affective state (2, 13).

A few experimental studies have shown neuronal modulations in non-human primates' prefrontal cortex, amygdala and ventral striatum related to self and other's actions and reward monitoring during monkey-monkey and monkey-human interactions (14–20), but the activity of the AI and the encoding of aversive stimuli during social interactions remains unexplored. We previously reported that macaque monkeys take into account their partners' welfare and display empathy during social decision-making involving both aversive and rewarding stimuli (21). Here we investigated neuronal activity in the AI associated

with aversive and appetitive outcomes for self and a monkey partner. Briefly, we show that AI neuronal activity describes experienced as well as observed outcomes, consistent with human imaging work on vicarious emotions. However, the majority of neurons showed preferential firing for self or other as recipient and only a minority showed mirror activity.

Two monkeys (*macaca fascicularis*) who had been tested in a prior behavioral study and identified as prosocial (M1 and M4 from Ballesta and Duhamel, 2015) participated in the present experiment. In order to investigate neuronal responses to different personal and vicarious outcome experiences, we used as a main task an imperative protocol in which monkeys had to manually press a key in response to a single cue associated with different outcomes for self, other or no one (as a control condition) (Fig. 1A). This ensured that the monkeys would sample the full range of stimuli, including those that, given a choice, they would rather avoid. The positive outcome was a small drop of liquid reward and the negative one a brief air puff near the corner of one eye. Monkeys reversed their role as actor and observer after completion of each block of 12 trials. As the actor was rewarded personally on only a fraction of the trials, we provided an additional incentive with a single big reward trial inserted at the end of each block. Active engagement in the task and role exchange guaranteed that the animals monitored the cues and outcomes concerning both self and the partner. Ocular and manual responses demonstrate that they tracked the task events both as actor and as observer, and that they understood the principle of turn taking (Fig. 1B-C and fig. S1). Conditions involving negative outcomes are of special interest. In the case of an air puff to self, eye blink rate began to rise about half a second before, then followed by a large blink response at the time of outcome delivery. In the case of air puff to the partner, we observed the same anticipatory and sharp albeit smaller blink responses (Fig. 1B-C), which can be considered as a form of emotional contagion or mimicry. Eye gaze data further confirm the monkeys' understanding of the visual cues. During outcome expectancy, fixation time in an area of interest (AOI) corresponding to the partner's face (a proxy of social attention) increased for all outcomes destined to self or to other, as opposed to a non-live agent (Fig. 1E). Finally, in a dual-choice task performed during the recording

session after completion of the imperative protocol, both animals exhibited a significant prosocial decision tendency (fig. S2).

We analyzed the activity of 148 single AI neurons (monkey 1: $n = 108$, monkey 2: $n = 40$). Half of the recorded cells (74/148, 50%) showed significant responses during one or several task epochs. In order to assess the information encoded in each cell's discharge, we tested the main effects of and interactions between the following factors: donor (self or other), recipient (self, other or nobody), and valence (reward or puff), using analyses of variance and multiple comparisons tests. A total of 32/148 (22%) and 14/148 (9%) cells responded at the time of cue appearance and cue selection, respectively. Despite displaying robust firing rates, most visually-responsive AI neurons were either non-selective (i.e. responded equally to all visual cues), or encoded the identity of the actor or of the outcome recipient. Furthermore, only a minority of these cells were selective to aversive outcomes (2 cue-related and 1 response-related cells) and therefore are not considered further. In contrast, AI neurons responding at the time of outcome delivery (39/148, 26%) were highly sensitive to valence and distinguished between reward and puff outcomes. We computed a simple valence selectivity index for each outcome-responsive neuron ($VI = [\text{Reward} - \text{Puff}] / [\text{Reward} + \text{Puff}]$). The distribution of VI for outcome-related activity (fig. S4) revealed two distinct functional subsets with little overlap between cells responding preferentially to pleasant (reward, $n = 16$) or unpleasant (puff, $n = 22$) stimuli, with one cell only responding indistinctly to both stimuli. Regarding donor effects, 3/39 (8%) cells responded to self as agent, 6/39 (15%) to other as agent, and 30/39 (77%) showed no significant difference between self and other as agent. Importantly, outcome-related neuronal activity depended on the identity of the recipient. Representative examples of single unit activity in response to unpleasant stimuli (Fig. 2A-C) show cells responding preferentially to a puff delivered to self, to other, or to both. The distribution of recipient selectivity indices ($RI = [\text{Self} - \text{Other}] / [\text{Self} + \text{Other}]$) was bimodal, indicating that most neurons responded preferentially to one of the two recipients and that few cells responded in a mirror-like manner to both experienced and observed aversive stimuli (Fig. 2D). The same analysis for cells encoding pleasant

stimuli showed a very different pattern, as we only found neurons encoding reward to self (Fig. 2E-F). In view of this striking dissociation, a further analysis was carried out on big reward trials, occurring at the end of each block. Because of its size and low frequency, this was a highly salient event. 28/148 (20%) cells responded to this particular outcome. As for small rewards, the RI distribution for big rewards is shifted toward own outcome encoding, but to a lesser degree, as a proportion of the neurons encoded other's reward only or both self and other's rewards (Fig. 2G-L).

The encoding of experienced and observed outcomes in the AI is summarized by the average population activity of all outcome-responsive neurons (51/148, 34%, i.e. all cells responding to either puff, small reward or big reward). It shows encoding of both monkeys as recipients for the air puff, of self only for the small reward, and a stronger encoding of self than other for the big reward (Fig. 3A-C). The topographic distribution of recording sites within the insula was examined using anatomical brain scans and electrode depth information. Outcome-selective cells were principally located in region of the anterior insula posterior to the limen, most likely including agranular and dysgranular sectors (22). The different cell classes were largely intermingled and no systematic spatial pattern reflecting the functional dimensions of valence or recipient coding could be identified (Fig. 1D). To sum up, joint coding of experienced and observed outcomes is demonstrated for negative and for highly salient positive experiences. Overlap between self and other representations is present at anatomical and functional levels, but the functional overlap is only in part attributable to the presence of neurons with mirroring properties and is mainly a reflection of the aggregate activity of subsets of cells responding preferentially to experienced or to observed outcomes.

A key issue about the encoding of observed outcomes is what drives such neuronal activity. Air puffs and liquid rewards generate multiple signals, such as sounds from the physical devices (e.g. the hiss of air flowing out, reward solenoid clicks) and visual cues (facial mimics). The presence of device noise is common to outcomes to self, other and nobody, and therefore cannot explain selective responses to other's outcome only, or to both own and other's but not nobody's outcome. However, air puffs and liquid rewards bring about behavioral reactions (e.g. eye blinks, mouth suction movements) that can lead

to enhanced attention toward the monkey partner. Therefore, it is crucial to determine whether neurons specifically encode the nature of the outcome and its affective/motivational value, or whether they respond whenever the partner's face enters the monkey's field of view. In a first step, we assessed neuronal responses to different gaze orientations during periods remote from the outcome epoch. None of the recorded neurons showed significant variations in firing rate while monkeys looked at the workspace, or during fixations on the partner's face keeping its gaze averted or making eye contact (Fig. 4A). Therefore, looking at the partner, even in a condition of high social attention engagement involving mutual gaze, does not by itself cause the neurons to respond. Yet it remains possible that these cells are tuned to specific facial cues that occur only at the time of the outcome. If that were the case, one would expect firing rates to depend on the monkeys' gaze direction at that particular instant. As monkeys were free to look toward or away from their partner's face we were able to compute, for each neuron, the number of eye fixation samples located inside the face AOI during the outcome delivery period (defined as 500 ms and 1000 ms from the onset of air puffs and big rewards, respectively). The trial set was split at the median of the fixation distribution in order to pit outcome-related activity for "gaze toward" (median percent face AOI fixations = 71.5%, iqr = 29.3%) against its activity for "gaze away" trials (median percent face AOI fixations = 18.5%, iqr = 22.3%). Results show that every neuron responded in both instances (Fig. 4B). However, most data points are located above the diagonal of the scatter plot, with the mean firing rates being on average 21% larger when gaze is directed to the recipient monkey's face than when directed away, a small but significant enhancement effect ($t_{26} = 2.42$, $P < 0.023$). Thus, although gaze on the partner's face is neither a necessary nor sufficient condition of the cells' response, facial cues could amplify neuronal encoding of others.

In this study, we show that neurons in the anterior insula (AI) respond to self and other's aversive and positive experiences. A subset of the recorded neurons responded to directly experienced unpleasant air puff stimuli, coherently with findings in humans showing insular properties in coding pain and negative emotions (23, 24). Another group of cells responded to liquid rewards, which is also consistent with

prior works showing that the AI receives input from adjacent primary gustatory cortex (25) and plays a role in reward expectation (26). Remarkably, we found neurons in the AI that responded only to observed negative outcomes. Gazing at a monkey partner receiving the aversive stimulus enhanced the activity of these cells, but observing the same monkey in a neutral context or while it was receiving a reward did not. Such a signal therefore appears ideally suited to represent the perceived emotional state of a conspecific experiencing an unpleasant event and ultimately facilitate prosocial behavior. Consistently with this hypothesis, electrical micro-stimulation of the same AI region in monkeys has been found to evoke affiliative lip smacking responses that, remarkably, occurred only in a social interaction context involving face-to-face contact between the experimenter and the animal (27). For positive outcomes, we did not find neurons encoding other's small rewards delivered in a similar context as the aversive air puff, but some cells responded to the other's big reward. A possible interpretation is that other's reward value must be larger in order to reach a level of saliency sufficient for generating a vicarious reward signal, i.e. the positive side of empathy (17). However, the other's big reward also cued the observer about the upcoming role reversal. The macaque anterior cingulate and lateral prefrontal cortex are interconnected with the AI and have recently been shown to play a role in the monitoring of self and other's actions during reward-based social decision paradigms (18). Thus, aside from encoding the hedonic value of other's rewards, the signal present in AI neurons might reflect enhanced arousal and motivation levels associated with being called into action.

The seminal discovery that the anterior insular cortex is activated by both physical and empathic pain has offered strong empirical support to the idea that the recruitment of own pain circuitry helps to understand the pain of others (28). Such overlapping activations are in fact sometimes used as an operational definition of empathy, with the tacit assumption that empathic feelings are mediated by specialized mirror neurons for self and other's pain. However, patients with congenital insensitivity to pain, who cannot rely on emotional resonance mechanisms, show normal AI activation during an empathy task (29). The failure to find a substantial number of neurons responding to both self and other's aversive outcomes appears as a further challenge to the mirror-matching account of insular function in empathy. Given that the estimated proportion of action-related mirror neurons in F5 area is

around 20% (30), it remains possible that our sampling was incomplete and not fully representative of the different cell subtypes present in AI. However, our population analyses indicate that the collective activity of neuronal pools encoding either self or other could account for overlapping activations reported in fMRI studies, even in absence of mirror neurons. Felt and observed emotions might share several features but nonetheless constitute distinct subjective experiences, ultimately calling for distinct behavioral responses. It has been argued that “true” empathy involves perspective taking and that this process requires an awareness of self as distinct from others (3). Some authors proposed that AI plays a key role in generating an integrated representation of the sentient self (31, 32). The capacity for self-awareness has long been denied to monkeys, but this view has been recently challenged (33), and the emerging evidence for distinct self and other representations at the neuronal level is calling for more in depth investigation of evolutionary continuities and discontinuities in social cognitive mechanisms among primates.

References

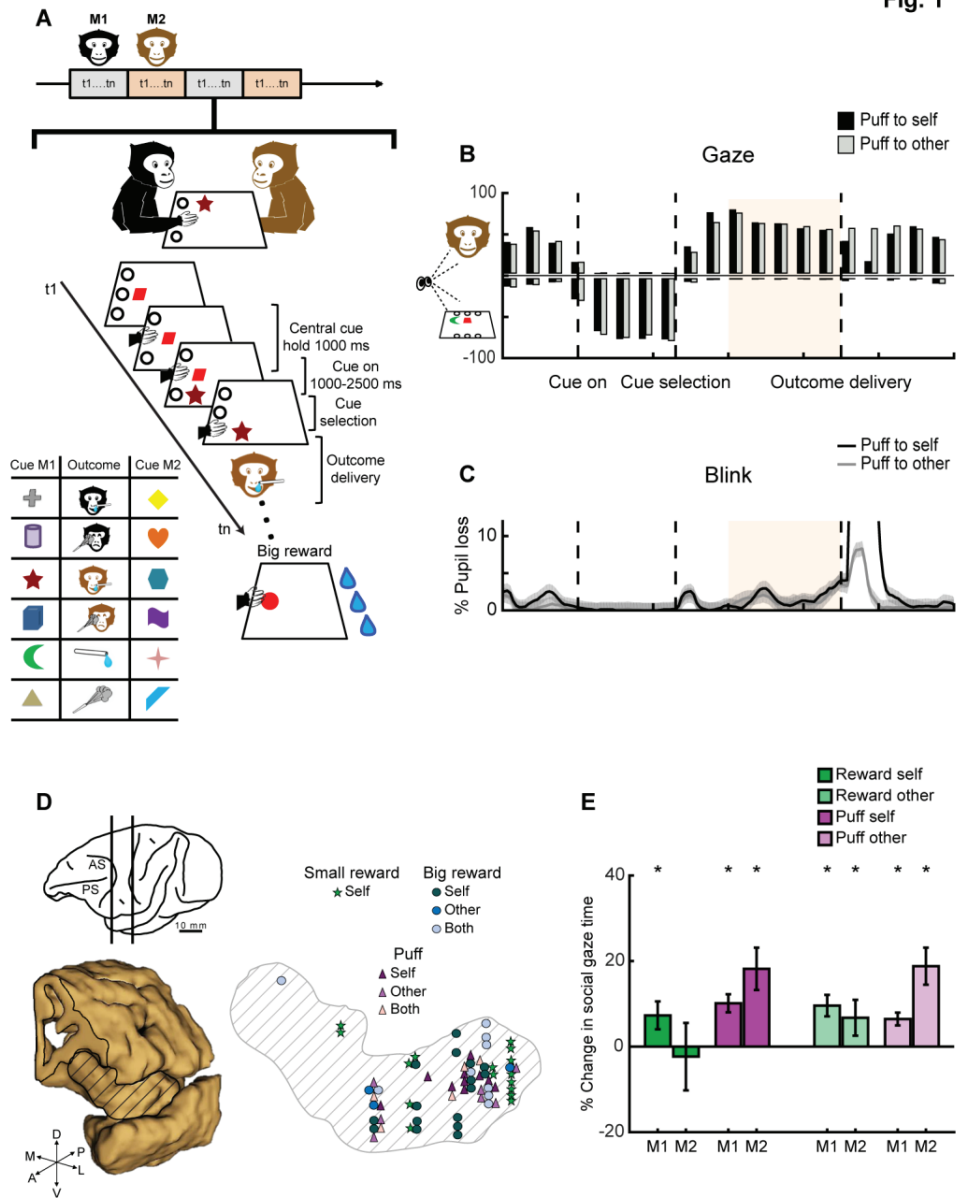
1. S. D. Preston, F. B. M. de Waal, Empathy: Its ultimate and proximate bases. *Behav. Brain Sci.* **25**, 1-20; discussion 20-71 (2002).
2. V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* **2**, 493–501 (1998).
3. C. D. Frith, U. Frith, Interacting minds--a biological basis. *Science.* **286**, 1692–1695 (1999).
4. J. H. Masserman, S. Wechkin, W. Terris, "ALTRUISTIC" BEHAVIOR IN RHESUS MONKEYS. *Am. J. Psychiatry.* **121**, 584–585 (1964).
5. D. Cheney, R. Seyfarth, B. Smuts, Social relationships and social cognition in nonhuman primates. *Science.* **234**, 1361–1366 (1986).
6. F. B. M. de Waal, Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* **59**, 279–300 (2008).
7. I. Ben-Ami Bartal, J. Decety, P. Mason, Empathy and pro-social behavior in rats. *Science.* **334**, 1427–1430 (2011).
8. J. P. Burkett *et al.*, Oxytocin-dependent consolation behavior in rodents. *Science.* **351**, 375–378 (2016).
9. R. J. Morecraft, K. S. Stilwell-Morecraft, J. Ge, P. B. Cipolloni, D. N. Pandya, Cytoarchitecture and cortical connections of the anterior insula and adjacent frontal motor fields in the rhesus monkey. *Brain Res. Bull.* **119**, 52–72 (2015).
10. T. Singer *et al.*, Empathy for pain involves the affective but not sensory components of pain. *Science.* **303**, 1157–1162 (2004).
11. G. di Pellegrino, L. Fadiga, L. Fogassi, V. Gallese, G. Rizzolatti, Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180 (1992).
12. G. Rizzolatti, L. Fogassi, V. Gallese, Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670 (2001).
13. M. Iacoboni, Imitation, empathy, and mirror neurons. *Annu. Rev. Psychol.* **60**, 653–670 (2009).
14. J. C. B. Azzi, A. Sirigu, J.-R. Duhamel, Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc. Natl. Acad. Sci.* **109**, 2126–2131 (2012).
15. K. Yoshida, N. Saito, A. Iriki, M. Isoda, Social error monitoring in macaque frontal cortex. *Nat. Neurosci.* **15**, 1307–1312 (2012).
16. S. W. C. Chang, J.-F. Gariépy, M. L. Platt, Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* **16**, 243–250 (2013).

17. S. W. C. Chang *et al.*, Neural mechanisms of social decision-making in the primate amygdala. *Proc. Natl. Acad. Sci.* **112**, 16012–16017 (2015).
18. K. Haroush, Z. M. Williams, Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell.* **160**, 1233–1245 (2015).
19. R. Báez-Mendoza, C. R. van Coeverden, W. Schultz, A neuronal reward inequity signal in primate striatum. *J. Neurophysiol.* **115**, 68–79 (2016).
20. R. Falcone, R. Cirillo, S. Ferraina, A. Genovesio, Neural activity in macaque medial frontal cortex represents others' choices. *Sci. Rep.* **7**, 12663 (2017).
21. S. Ballesta, J.-R. Duhamel, Rudimentary empathy in macaques' social decision-making. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 15516 (2015).
22. S. T. Carmichael, J. L. Price, Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J. Comp. Neurol.* **346**, 366–402 (1994).
23. R. C. Coghill, C. N. Sang, J. M. Maisog, M. J. Iadarola, Pain intensity processing within the human brain: a bilateral, distributed mechanism. *J. Neurophysiol.* **82**, 1934–1943 (1999).
24. L. J. Chang, P. J. Gianaros, S. B. Manuck, A. Krishnan, T. D. Wager, A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. *PLoS Biol.* **13**, e1002180 (2015).
25. E. T. Rolls, Functions of the anterior insula in taste, autonomic, and related functions. *Brain Cogn.* **110**, 4–19 (2016).
26. T. Mizuhiki, B. J. Richmond, M. Shidara, Encoding of reward expectation by monkey anterior insular neurons. *J. Neurophysiol.* **107**, 2996–3007 (2012).
27. F. Caruana, A. Jezzini, B. Sbriscia-Fioretti, G. Rizzolatti, V. Gallese, Emotional and Social Behaviors Elicited by Electrical Stimulation of the Insula in the Macaque Monkey. *Curr. Biol.* **21**, 195–199 (2011).
28. J. Zaki, T. D. Wager, T. Singer, C. Keysers, V. Gazzola, The Anatomy of Suffering: Understanding the Relationship between Nociceptive and Empathic Pain. *Trends Cogn. Sci.* **20**, 249–259 (2016).
29. N. Danziger, I. Faillenot, R. Peyron, Can we share a pain we never felt? Neural correlates of empathy in patients with congenital insensitivity to pain. *Neuron.* **61**, 203–212 (2009).
30. M. Maranesi *et al.*, Anatomic-functional organization of the ventral primary motor and premotor cortex in the macaque monkey. *Eur. J. Neurosci.* **36**, 3376–3387 (2012).
31. A. D. B. Craig, How do you feel--now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* **10**, 59–70 (2009).
32. H. D. Critchley, S. Wiens, P. Rotshtein, A. Ohman, R. J. Dolan, Neural systems supporting interoceptive awareness. *Nat. Neurosci.* **7**, 189–195 (2004).
33. L. Chang, S. Zhang, M.-M. Poo, N. Gong, Spontaneous expression of mirror self-recognition in monkeys after learning precise visual-proprioceptive association for mirror images. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 3258–3263 (2017).

Acknowledgments

The authors of this article are grateful to Serge Pinède for technical support, Jean-Luc Charieau and Fidji Francioli for expert animal care, and Chloé Dupuis for assistance in recording site localization. Pier Francesco Ferrari provided helpful advice on earlier versions of the manuscript. This work was supported by the LABEX (ANR-11—LABEX-0042) of University de Lyon within the program “Investissement d’Avenir” and by grants from the Rhône-Alpes Region and from the Agence Nationale de la Recherche (BLAN-SVSE4-023-01, BS4-0010-01) to JRD.

Fig. 1



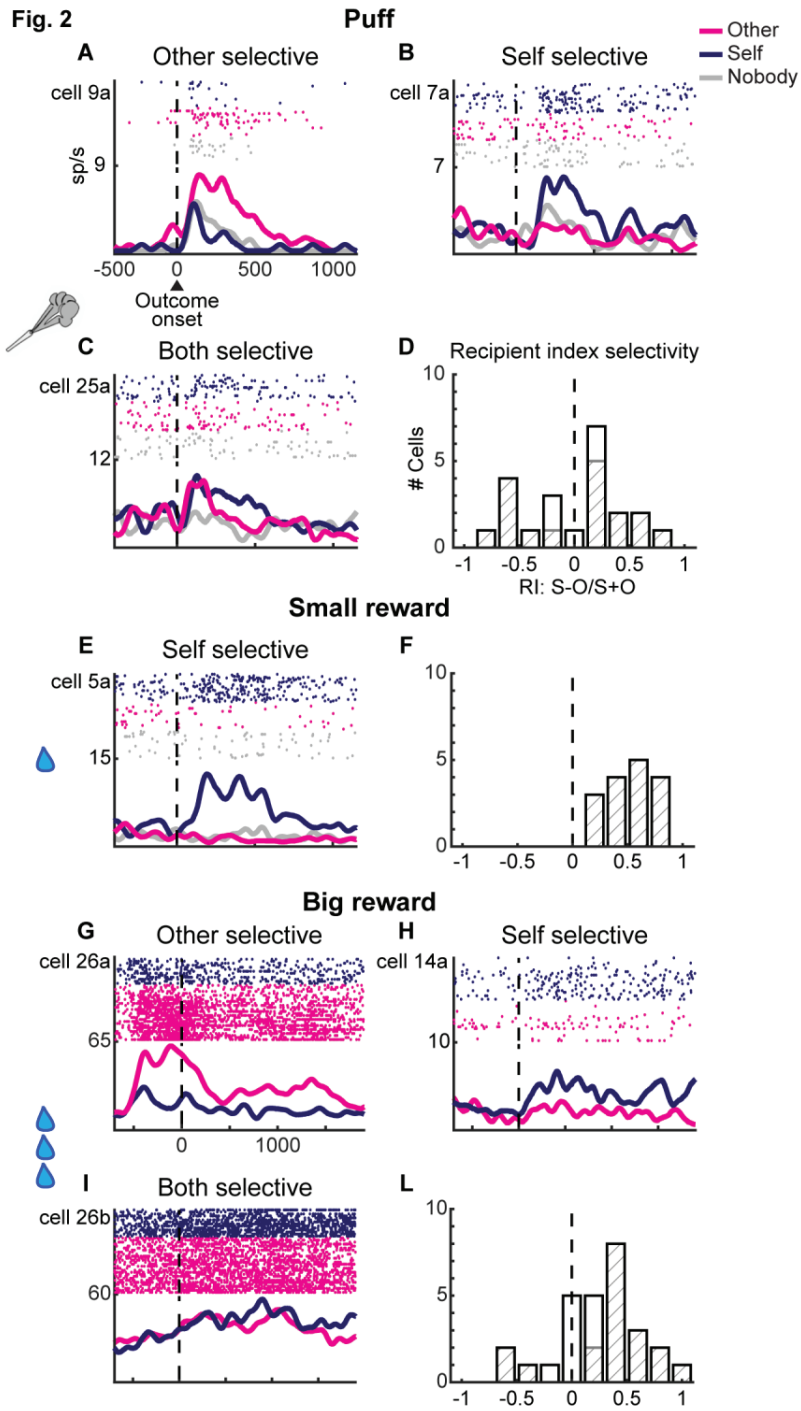


Fig. 3

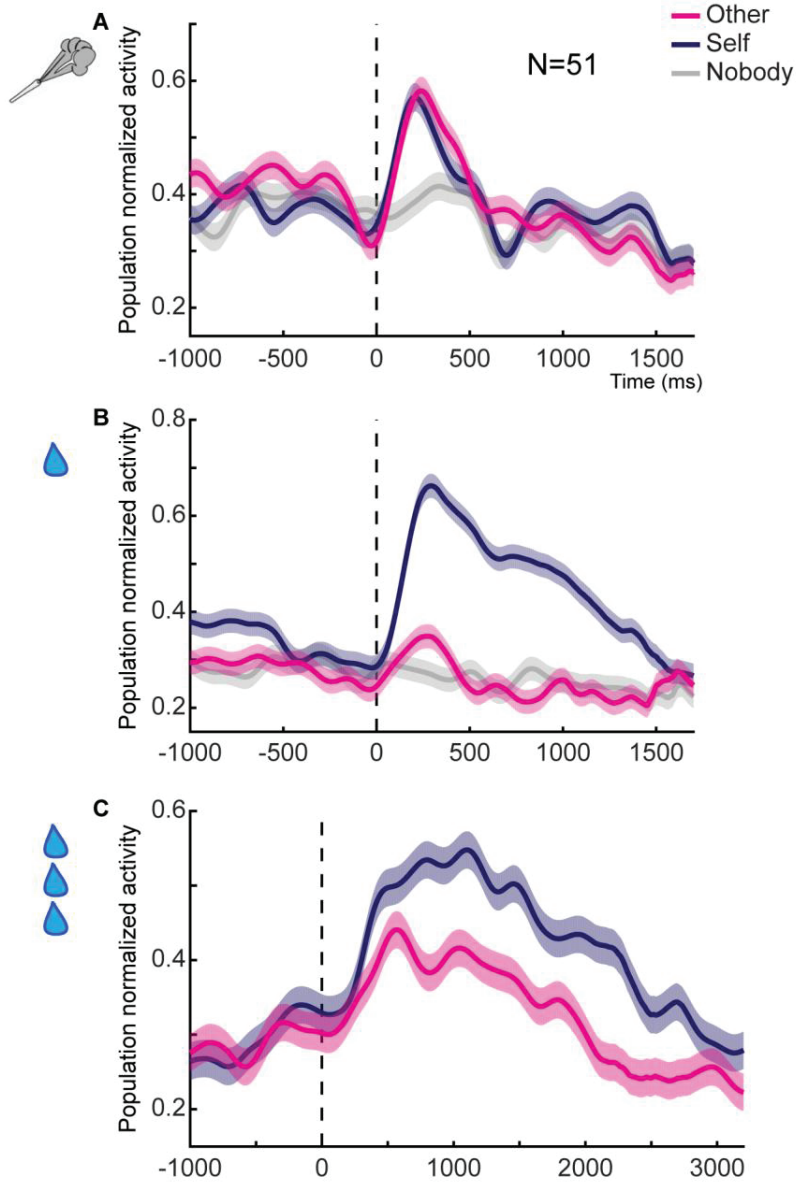


Fig. 4

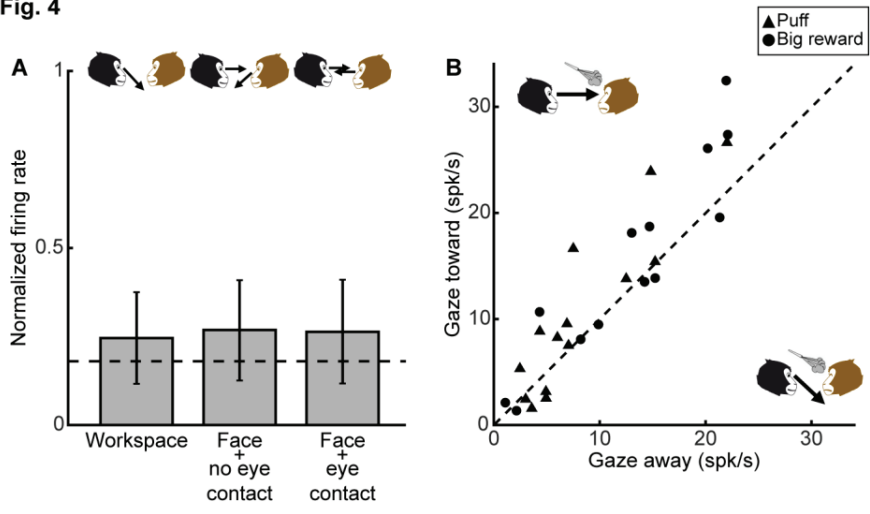


Figure legends

Fig. 1. Task, behavior and outcome cells brain distribution. A) Temporal sequence of events of one example trial of the imperative task. The table contains the possible different cues for monkey 1 and monkey 2 related to the different kinds of outcome. B-C) Oculomotor behavior of monkey M1 as actor during puff to self (black) and puff to other (grey) trials (mean of $n = 25$ sessions). B) Percentage of gaze fixations toward face or workspace. C) Eye blink rate expressed as percentage of pupil signal loss. Same alignment as (B). D) Topographical distribution of outcome responsive cells for small reward, big reward, and puff for self, other, and both recipients in the recorded insular portion. E) Percentage increase and decrease in social gaze time during experienced and observed outcomes expectancy (background shading in B-C) for liquid reward (green bars) and puff (purple bars) to self and other ($* P < 0.01$). The baseline represents gaze time toward the non-live recipient.

Fig. 2. Recipient-selective outcome responsive cells. A) Raster and spike density plots of a cell responding significantly to other's puff (magenta) compared to self (blue) and nobody (grey) puff. Each dot in the raster plot represents the cell discharge relative to the outcome onset. B) Raster plot of a cell responding selectively to self puff. C) Raster plot of a cell responding selectively to both monkeys' puff. D) Distribution of recipient selectivity index computed for all puff-responsive cells. E) Raster plot of a cell responding to self small reward. F) Recipient selectivity index computed for all small reward-selective cells. G) Raster plot of a cell responding to other's big reward. H) Raster plot of a cell responding to self big reward. I) Raster plot of a cell responding to both monkeys' big reward. L) Recipient selectivity index computed for all big reward-selective cells. The raster plots conventions are the same as those of Fig. 2A. Dashed portions of bars in D-F-L indicate proportion of cells with significant preference for self or other as recipients.

Fig. 3. Population activity of outcome cells to puff, small, and big reward. A) Population normalized activity computed for all outcome-related cells ($n = 51$) in response to air puffs to self, other or nobody. Magenta and blue curves represent live recipients (other - self), grey curves indicated non-live recipient (nobody). Activity is aligned on cells' response latency. Shaded error bars represent \pm SEM. B) Population normalized activity for the small reward. Same alignment of (A). C) Population normalized activity for the big reward. The duration of big reward delivery was longer than the other outcomes (2000 ms).

Fig. 4. Neuronal activity modulation during workspace or other's face fixations. A) Mean normalized activity of cells encoding other's outcome (with or without mirroring responses, $n = 27$) during fixation on the workspace and on the other's face without and with presence of eye contact. Cue presentation and outcome epochs were excluded from this analysis. Dashed line represents 'baseline' activity measured during fixations outside of the two AOIs. B) Scatter plot showing, for the same cell subpopulation in A, outcome-related activity associated with low *versus* high amount of face fixation during outcome delivery. Black triangles and circles represent puff and big reward cells, respectively.

Supplementary

Materials and Methods

Animals

This study involved experiments on non-human primates. All surgical and experimental procedures were in conformity with current guidelines and regulations on the care and use of laboratory animals (European Community Council Directive No. 86–609) and were conducted in authorized facilities (Department of Veterinary Services, Health & Protection of Animals, permit number 69 029 0401). The specific research protocol was examined by CELYNE, the local ethics board, which approved the in vivo methods used in this experiment (authorization No. 2015061213048343). Three male long-tailed macaques (*Macaca fascicularis*) (monkey 1, M1, aged 8 years, weight 7.5kg; monkey 2, M2, aged 7.5 years, weight 7kg; monkey 3, M3, aged 7.5 years, weight 8kg) participated in the experiment. Electrophysiological recordings were conducted in M1 and M2 only. M1 was the dominant monkey and M2 and M3 were subordinates. Animals were initially housed as a mini-colony in a large enclosure where they could have direct physical interaction, while during the period of the experiments they were isolated in individual cages. When isolated, monkeys could still communicate vocally and visually thanks to the combination of glass and wire mesh dividers between cages. Home cages were enriched with different toys and substrate on the floor to allow foraging. Animals were fed with monkey chow, fresh and dry fruits, vegetables. Water intake was regulated during training and data collection periods since the behavioral task involved fluid rewards, but they were allowed one day of free access to water each week.

Surgery

Monkeys underwent two sterile surgical interventions under isoflurane anesthesia (1.0 - 3.0%). Pre-anesthesia was induced with intramuscular injections of ketamine (10mg/kg) and of medetomidine (25µg/kg). Heart and breathing rate, blood pressure, oxygen and CO2 levels, and body temperature of

the animals were monitored continuously. During the first surgery, we implanted an MRI compatible head-restraint device serving to immobilize monkeys' head during the experiment. During the second surgery, a recording chamber was implanted over a craniotomy performed on the left hemisphere (M1) or on the right hemisphere (M2). The positioning of the recording chamber was calculated using stereotaxic coordinates derived from an MRI scan performed under tiletamine/zolazepam (15mg/kg) anesthesia between the two surgeries and confirmed with a second scan performed in the weeks after the second surgery. Animals were allowed at least one month to recover and received antibiotic treatments and adapted pain management as needed.

Apparatus

Experiments were conducted in a semi-dark room where two monkeys were seated face to face in a primate chair. The distance between the animals allowed them to see each other but not to interact physically. The working space positioned between the two animals consisted in a white horizontal board (30cm X 45cm) on which two sets of three response keys were mounted. Computer generated cues were projected near the response keys by means of a video projector positioned above the board. Monkeys could respond by pressing the corresponding keys in order to initiate a trial and a given outcome, as explained below. Each monkey had a feeder tube placed near its lips to deliver drops of water (200 ms, ~ 1 ml), using a gravity-based solenoid device. Discrete air puffs (4 bars, 60 ms duration) could also be delivered close to the monkeys' left or right eye through a tubing system connected to solenoid device and pressure gauge. Additional solenoid devices placed on the side of the workspace allowed to deliver control rewards and air puffs to "nobody". Each animal's eye position was monitored using Eye-Trac 6 (ASL, Bedford MA) infrared video eye trackers (200-Hz sampling rate). At the beginning of each session the eye tracker signals were calibrated by using a regular grid of 3x3 LEDs mounted on opaque panel inserted vertically at 40 cm of distance from monkey's faces. The monkeys were rewarded for fixating each point in turn and the procedure was repeated until satisfactory calibration was obtained for

both animals, after which the board was removed. Behavioral control and visual stimulus presentation was achieved using the REX/VEX system (34).

Behavioral procedures

The general task principle was based on the social decision-making paradigm used by Ballesta and Duhamel (2015), adapted for electrophysiological recordings. It included as a main task an imperative protocol, in which monkeys had to actively select a single option on each trial, and a simplified choice protocol designed for the purpose of assessing the monkeys' preferences among the different options used in the imperative protocol. We opted for such a task design because our goal was to investigate neuronal responses to different personal and vicarious outcome experiences. In a decision-making task in which the participant exhibits systematic choice tendencies, some options may rarely or never be chosen. The imperative paradigm thus ensured that monkeys would sample the full range of stimuli, including those that, given a choice, it would have avoided. Such an operant task was preferred over a more simple classical conditioning paradigm because it kept the subjects actively engaged in monitoring cues and outcomes, an important requirement when monkeys are personally concerned as outcome recipient in only a fraction of the trials.

Imperative protocol

This task was divided in multiple short blocks with the two monkeys alternating in their role as actor and observer. The first trial of each block, and all subsequent trials within a block, began with the illumination of a central cue on the side of the workspace of the designated actor monkey. The monkey had to press and hold the central response key in order to initiate the trial sequence. Shortly thereafter (1000 ms), a single cue appeared in one of the lateral positions selected from a randomized schedule. The monkey had to wait for a variable delay (1000 - 2500 ms) until the extinction of the central cue (“go signal”) in order to press the key corresponding to the lateral cue. If the monkey responded within 1000

ms, the outcome associated with the lateral cue was delivered after a delay of 2200 ms. Successful completion of the ongoing trial conditioned progression to the next trial. If the monkey released the central key too early or failed to select the lateral key, the same cue was presented until the monkey succeeded. Six different cue-outcome associations were used, and each one was presented twice in a given block, for a total of 12 trials. The six possible outcomes, illustrated in Fig. 1A (table), were: (1) reward to self, (2) air puff to self, (3) reward to other, (4) air puff to other, (5) reward to nobody (a drop of water dropped in a container placed on one side of the workspace), (6) air puff to nobody (a jet of air in empty space generating noise but no tactile stimulus). Because under such an outcome schedule, the actor monkey was personally rewarded on only 16.6% of the trials, we needed to provide it additional incentives to complete the task. Each block therefore ended with a single “big reward” trial, instructed by a specific central cue, which when selected provided 2000 ms of free access to water (~ 10 ml). This last trial also signaled the role switching of the monkeys as actor and observer. Each outcome type was associated with a unique cue, and each monkey had its own set of images. Images were renewed every week. The monkey actors completed a minimum of 10 blocks of imperative trials, for a total of 120 trials.

Choice protocol

The choice protocol followed completion of the imperative protocol and differed from the latter by the use of two lateral cues and the requirement to select one of the two options upon extinction of the central cue. The proposed decisions and choice preferences of the two monkeys used for recordings were: (1) reward to self v. puff to self, (2) reward to other v. puff to other (fig. S2, table). The spatial position of the two cues was randomly selected on each trial and the timings, delays and turn taking procedure were identical to those used in the imperative protocol.

Electrophysiological recordings

Neuronal recording were conducted over a period of approximately 8 months. M1-M2, M1-M3 and M2-M3 dyads were tested, with a majority of recordings sessions obtained with the M1-M2 dyad. We generally recorded from one monkey at a time, although some sessions involved simultaneous recording in the two monkeys. Neuronal signals were acquired using 16 channels U-probe (Plexon, Dallas) tungsten electrodes (15 μ m diameter, impedance 275 \pm 50 k Ω) with an inter-contact spacing of 300 μ m. Electrodes were connected via a ribbon cable to the Plexon PBX preamplifier (gain of x1000), and from the amplifier to a Spike2 (CED, Cambridge) data acquisition system. Electrical brain signals were sampled and stored at a frequency of 20KH, Electrodes were lowered into the targeted brain tissue using a motorized multi-electrode drive (NAN Instruments, Nazareth) at to a recording chamber and grid system (Unimécanique, Pontoise). Single unit spike waveforms were presorted online and resorted offline using Plexon Offline Sorter software. Sorted spike channels were then merged with eye position data and task event information for subsequent analyses using custom software developed in the Matlab environment (The MathWorks Inc., Natick, MA).

The monkeys have not been euthanized at the end of the experiment. We reconstructed the approximate location of the recording sites thanks to an anatomical brain scan performed with an MRI-compatible recording grid filled with iodopovidone as contrast agent. This provided a reference frame for calculating the origin and angle of the electrode trajectory of each of the grid location. By combining this data with recorded information of electrode depth we inferred the mostly likely location of each electrode contact in the brain volume. We then used ITK-SNAP (University of Pennsylvania, PA) software to virtually “cut” a block of tissue, rotate this volume and partially open the lateral sulcus in order to visualize the insular subregion containing our recording sites.

Data Analysis

Behavior

Subjective evaluation of the cue-outcome associations was assessed by simple calculation of choice preferences of each monkey in the choice protocol (fig. S2). We also examined two oculomotor variables previously shown to be sensitive behavioral markers during empathy-based decision making (21) : social gaze and eye blink rate. Social gaze was estimated from the recorded horizontal and vertical eye position data. For each monkey, we calculated the projected position of the partner's face and of the workspace into the visual field and defined these two surfaces as our main areas of interest (AOI). Eye position traces were segmented into discrete saccade and fixations using a standard algorithm (adapted from ASL File Analysis Tool, version 1.03). Each fixation episode was then labeled as belonging to the face AOI, the workspace AOI, or elsewhere. Eye blinks were calculated from the "pupil loss" signal generated by the eye tracker. Blink start and end events were scored if the pupil was lost or recovered, respectively, for at least 4 consecutive 5 ms-long records.

In order to characterize the behavior of the monkeys in their role as actor and observer and as a function of the expected outcome we performed statistical analyses on 25 representative experimental sessions. For illustration purposes, fig. S1 shows AOI fixations and eye blinks (along with recorded key press and release actions) averaged across sessions for all "puff self" trials of M1. Fig. 1 B-C show AOI fixations and eye blinks for all "puff self" and "puff other" trials of M1. We quantified the effects of the expected puffs and rewards on gaze and eye blink behavior by considering a 1500 ms window preceding outcome delivery. For each outcome type, a difference score was calculated between total social gaze time (i.e. fixation inside the face AOI) during this outcome anticipation epoch and a 'baseline' window of equal duration set in the intertrial interval. Six difference scores were obtained for each session and the magnitude of social gaze enhancement or suppression across all sessions was evaluated using a Wilcoxon signed rank test for zero median. The same procedure was used to evaluate eye blink rate enhancement and suppression.

Spike data

The activity of each recorded neuron was visually inspected by plotting raster and spike density functions aligned on the main task events. We searched for significant changes in firing rate using non-parametric Mann-Whitney U tests between spike counts in a pre-cue window (-500:0 ms from cue onset) and several windows centered on events of interest: a) cue appearance (0:500ms from cue onset), b) cue selection (0:700ms from manual response), c) delay (-900:0 from outcome delivery), d) outcome (0:700ms from outcome delivery). These windows were optimal for a majority of neurons, but could be individually adjusted when needed to take into account the neuronal response dynamics (e.g., highly phasic or highly tonic responses). As the task was constructed following a fully factorial design with Donor (self, other), Recipient (self, other, nobody) and Valence (air puff, reward) as independent variables, we used three-way ANOVAs to search for main effects and interactions of the three main factors during each of the considered epochs. Post-hoc pair-wise multiple comparisons (Tukey-Kramer method) were applied where justified. Separate analyses were performed on “big reward” trials, for which we used t-tests to assess the presence of recipient effects (self *versus* other) during the reward delivery epoch (0:1000ms from reward onset). The result of these analyses served as a basis for classifying cells in different subtypes.

We also computed, for each neuron, a set of selectivity indices that served to summarize donor, recipient and valence effects using a simple $[A-B]/[A+B]$ ratio and examined the distribution of these effects at the population level.

Finally, population activity plots of recipient effects were generated by computing average population spike density functions. Individual cell’s spike densities were normalized prior to averaging using the following normalization rule:

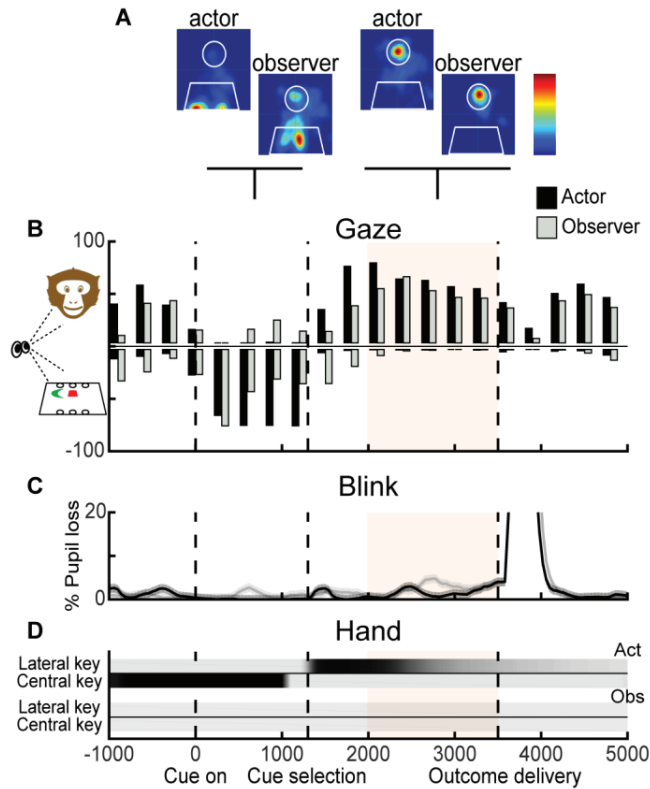
$$[\text{FiringRate} - \text{MinFiringRate}] / [\text{MaxFiringRate} - \text{MinFiringRate}]$$

Additional population analyses were performed on a subset of neurons in order to assess the influence of social gaze on their activity (Fig. 4A-B). The first analysis tested for the cells’ responses to gaze direction, independently of the occurrence of cues, manual responses, rewards or air puffs. The data used for this analysis were all of the fixation samples collected outside of those the critical epochs when

stimuli were presented on the workspace or outcomes delivered to self or to the other monkey. For each neuron, spike counts were calculated during all fixations lasting at least 300 ms and located inside the workspace or the face AOIs. For face fixations, we separated instances where the partner's gaze was averted when the eye landed on the face and remained averted (face, no eye contact), and instance where the partner was already gazing in its own face AOI, thus establishing eye contact (face, eye contact). Differences in neuronal activity between three gaze conditions were tested with a simple one-way ANOVA. The second analysis examined the modulation of outcome-related responses by gaze direction in order to determine if these responses depended on viewing the actual delivery of the outcome to the partner monkey. For each instance of reward or air puff delivered to the partner monkey, we determined the percentage of time spent looking in the face AOI during a period comprised between 0 and 500 ms from outcome onset. Spike discharges rate were then compared between low ($<$ median) and high (\geq median) face gaze trials. Across the cell population, "low face gaze" samples had a mean of 18.8% (std = 13.8%) and "high face gaze" had a mean of 68.7% (std = 17.6%) face gaze time. Differences in activity level associated with low and high face gaze during outcome delivery at the population level were tested with a simple two-sided t-test.

Supplementary Figures

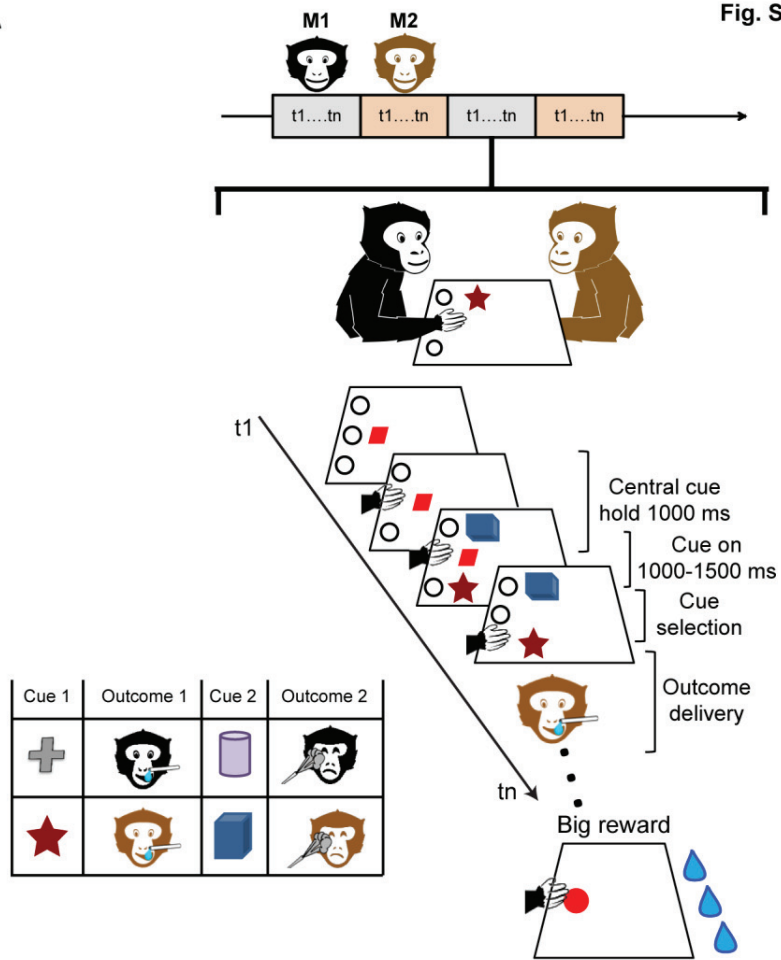
Fig. S1



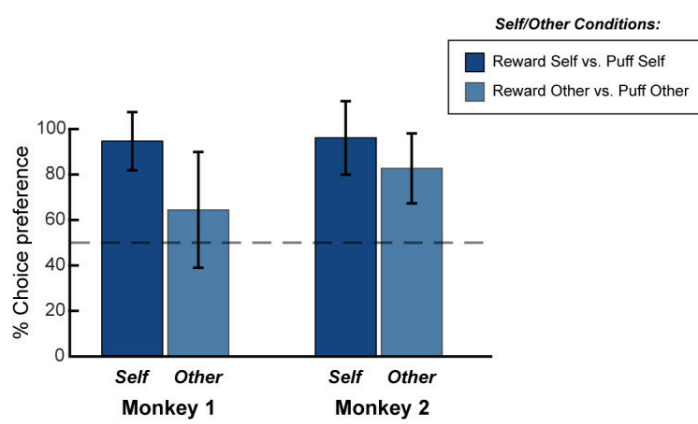
Oculomotor behavior during aversive outcome expectancy. A) Example fixation heat maps for the same monkey (M1) as actor or observer during cue presentation and outcome expectancy. Light outlines indicate workspace and face areas of interest. B-C-D) Oculomotor and manual behaviors of the same monkey as actor or observer during self puff trials (mean of $n = 25$ sessions). B) Percentage of gaze fixations toward face or workspace. C) Eye blink rate expressed as percentage of pupil signal loss. Same alignment as (B). D) Hand location on the workspace. Dark shading indicates high probability of the hand being on the central or peripheral response key. Note the absence of key presses of the monkey as observer. Same alignment as (B).

A

Fig. S2

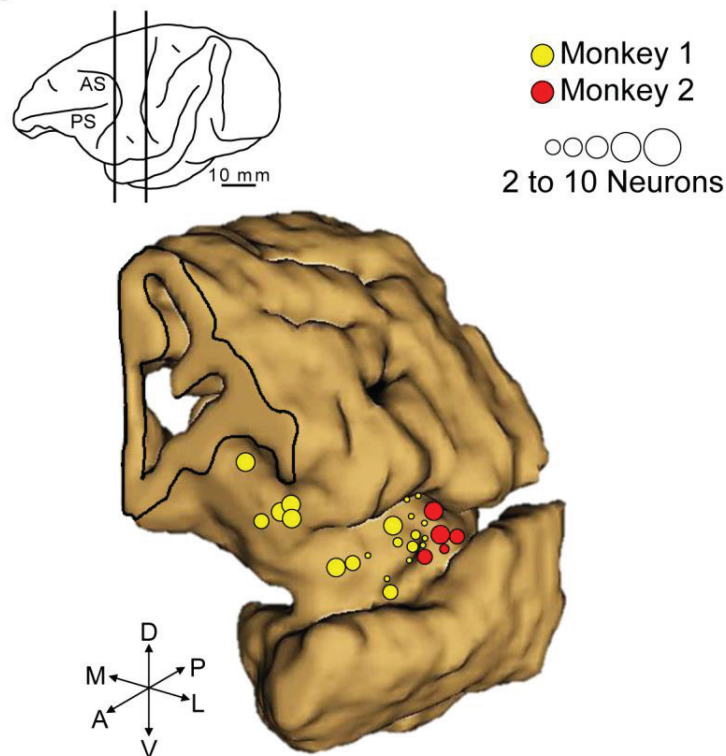


B



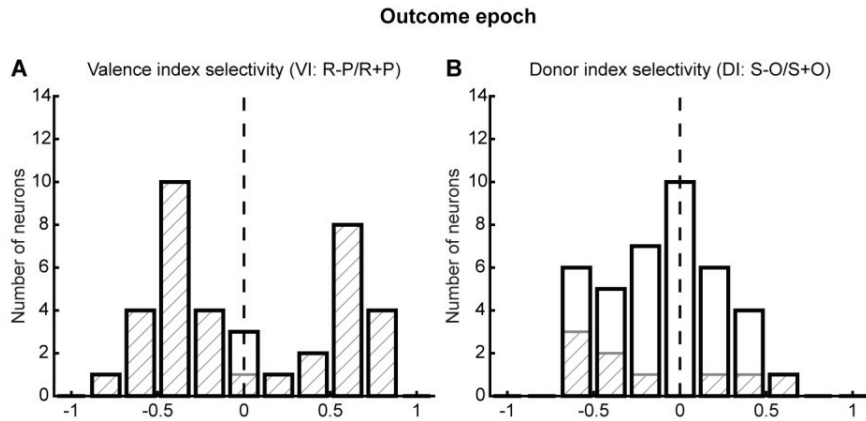
Choice task and behavior. A) Temporal sequence of events of one example trial of the choice task. The table represents the two possible cue combinations of the choice conditions for self and other as recipients. B) Percentage of choice preferences for self and other conditions for monkey 1 and monkey 2 on the total of sessions ($n = 25$). Dashed line represents the prosociality threshold fixed at chance level (50%). Both monkeys were significantly prosocial.

Fig. S3



Recording sites. Topographical distribution of recording sites for monkeys M1 (yellow circles) and M2 (red circles) in the targeted insular portion. Cells recorded in M1 spanned most of the AI region whilst those recorded in M2 were all localized in the portion of AI posterior to the limen. It should be noted however that the same types of neuronal selectivity for aversive and appetitive stimuli and for self and other as outcome recipients were present in both animals.

Fig. S4



Valence and Donor index selectivity. A) Valence sensitivity index distribution for all outcome responsive cells. Index computed during the outcome epoch of the task. The near absence of overlap between the puff and reward selective subpopulations justified analyzing these cells separately. Dashed portions of bars indicate significant cells. B) Donor sensitivity index distribution for all outcome responsive cells. Most outcome-related cells were not selective to the identity of the donor and data for self and other as donor were combined. For cells showing significant donor-selective activity, the analyses of valence and recipient effects were performed on trials associated with the preferred donor.

References

21. S. Ballesta, J.-R. Duhamel, Rudimentary empathy in macaques' social decision-making. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 15516 (2015).
34. A. V. Jr. Hays, B. J. Richmond, L. M. A. Optican, *A UNIX- based multiple process system for real-time data acquisition and control* **2**, (1982).

PART III

DISCUSSION

In our study, we engaged dyads of macaques to perform a social task where they could be, in turn, actors of performed trials and recipients of possible positive or negative outcomes. Our behavioral results were coherent with those obtained by Ballesta and colleagues (2015) who showed that monkeys considered the welfare of their conspecifics by making mostly prosocial decisions, such as rewarding the partner or avoiding the partner to receive an airpuff. In their work, the authors went further than the demonstration of monkeys' prosocial behavior towards each other and suggested evidences of empathy since they observed that animals were "reacting" significantly to the unpleasant event experienced by the other.

Following the same principles used in Ballesta and colleagues' paradigm, we designed a behavioral task adapted to electrophysiological recordings. In our study, monkeys were performing two blocks of trials: an imperative section followed by a choice block. Thanks to the imperative trials we could assess the neural responses to all types of outcomes. This was not possible by having only a choice task. Indeed, because of their rational and prosocial way of behaving, monkeys rarely chose to punish themselves with an airpuff instead of getting some reward, or to punish the partner instead of rewarding him. However, we used a choice task to verify the degree of pro-sociality of the animals, and we saw that monkeys' choices for each other were mostly prosocial. Moreover, the gaze behavior of the monkeys in the imperative trials revealed that animals discriminated the different cue/outcome associations and monitored their partners' actions and states. For example, during the time occurring between the choice and outcome delivery (delay epoch), monkeys' eye fixation to the partner's face (social gaze) was enhanced when a reward was about to be delivered to the partner. Conversely, social gaze decreased and eye blink rate augmented when an airpuff was about to be delivered to the observed monkey.

For our neural recordings we targeted the anterior insula (AI), one of the brain regions thought to be a neural marker of empathy. Numerous fMRI studies in humans showed how this area activated during both experienced and observed pain (Singer et al., 2004; Jackson et al., 2005; Lamm et al., 2007; Hein et al., 2010). The AI has been explored with other techniques as well, such as intracortical micro stimulation (ICMS) in primates. Indeed,

several studies in monkeys used ICMS of single cells in specific sectors of the insular cortex eliciting different behaviors. More explicitly, electrical stimulation of neurons in the anterior sector of insula evoked disgust related and ingestive behaviors, while stimulation of the dorso-medial sector of insula provoked forelimb movements, and finally stimulation to the mid-ventral section of insula produced lip smacking behavior (Caruana et al., 2011; Jezzini et al., 2012). Other studies have been conducted in human epileptic patients. These clinical works used intracranial stimulations and showed insular properties in processing both painful and non-painful somesthetic inputs, facial emotional expressions such as disgust, and ictal vomiting (Ostrowsky et al., 2002; Krolak-Salmon et al., 2003; Catenoix et al., 2008).

The described neurophysiological works in primates showed the neural substrates underlying different degrees of social behavior in monkeys. Several primates electrophysiological studies showed the modulation of the neuronal activity in the orbitofrontal cortex (OFC) by social context and social stimuli (Azzi and Duhamel, 2012; Watson and Platt, 2012). In other studies involving social interaction tasks it has been shown that single neurons in the ventral striatum of primates encode social actions related to own rewards and reward inequity (Bàez-Mendoza et al., 2016). Other works showed how single neurons in the frontal cortex are involved in the coding of social agency (of conspecifics and not) and in the monitoring of other's mistake (Yoshida et al., 2011, 2012, Falcone et al., 2017). In another study it has been reported that lateral prefrontal cortex is engaged in the computation of the sense of social competition and its relative result (Hosokawa and Watanabe, 2012). The coding of unknown social actions has been investigated in the dorsal section of the anterior cingulate cortex (dACC; Haroush and Williams, 2015).

Interestingly, among the cortical regions the ACCg has been shown as a brain area involved in the computation of social reward (Chang et al., 2013). Chang and colleagues (2013) reported that neurons in the ACCg respond to self and other's reward with a mirror-like pattern, in a consistent way with previous studies linking mirror-like properties to specific social functions like shared experience and empathy (Amodio and Firth, 2006). Among the limbic regions, amygdala has emerged as a critical neural marker in regulating social

decisions (Chang et al., 2015). Based on their findings, the authors claimed that neurons in the basolateral section of the amygdala reflected monkeys' tendency to make prosocial decisions for their partners and signaled the value of rewards for self and other with a mirror-like pattern.

In our task monkeys could experience three types of outcome: small reward, big reward, and airpuff. Our neural results show that neurons in the AI respond to self and other's aversive and positive experiences.

A subpopulation of neurons responded to directly experienced air-puff, coherently with previous findings on humans showing insular properties in coding aversive experienced events (Singer et al., 2004). Another group of cells responded to own small and big rewards, in a consistent way with a previous neurophysiological work showing that neurons in the AI play a role in reward processing (Asahi et al., 2006).

Remarkably, we found that neurons in the AI replied to observed negative outcome of the other monkey. Our analysis showed that neural activations to the partner's air-puff were not enhanced by the other's facial expressions during the reception of the aversive outcome, such as eye blink. Thus, we believe that these responses may reflect the internal emotional state of the other experiencing an unpleasant situation. For positive outcomes, we did not find neurons responding to other's small reward, but notably we observed a class of cells coding the other's big reward. This result is consistent with previous reports highlighting the insular properties in salience processing (Uddin, 2014).

In human's fMRI studies the AI emerged as the overlapping area activated for both experienced and observed pain (Wicker et al., 2003; Singer et al., 2004; Morrison et al., 2004; Botvinick et al., 2005; Jabbi et al., 2007; Lamm et al., 2007; Ochsner et al., 2008). Thus, it has been suggested that empathy for pain may involve a mirror-matching model of the affective and sensory features of others' pain. The theory of "embodied simulation" sustains that such overlap reflects an automatic resonance to other's affective states, leading to affect sharing and empathy (Gallese et al., 2004, 2007; Keysers and Gazzola, 2006). In addition to this "mirror-matching" mechanism, it has been suggested that perspective-taking processes provide the understanding of other's emotional states in a more reflexive

way (Decety and Jackson, 2004; Beer and Ochsner, 2006; Mitchell, 2006). Following this logic, it has been hypothesized that an observer who cannot feel a certain feeling might not be able to directly empathize with someone experiencing this feeling and would necessarily have to engage in a perspective-taking posture to understand the other's state (Singer, 2006). Patients with the rare syndrome of congenital insensitivity to pain (CIP) offered a unique opportunity to test this model of empathy. Indeed, clinical researches have explored how the lack of self-pain representation might influence the perception of other's pain (Dazinger et al., 2006, 2009). The reported results showed how in these patients the perception of other's pain was mainly computed in the anterior ventromedial prefrontal cortex (vmPFC) and in the ventral posterior cingulate cortex (vPCC). Therefore, these findings underlined that understanding someone else's feeling is possible despite the impossibility to have direct personal experiences of the perceived feeling.

Coherently, our findings sustain the idea that a mirror neural model is not sufficient to explain empathy for pain because it does not implicate a distinction of self and other's emotional internal states. Moreover, several works have well distinguished empathy from other similar concepts, such as emotional contagion and compassion, where there is a lack of self-other distinction (Singer and Lamm, 2009; Bird and Viding, 2014).

Taken together, our results extend the idea of a neural network processing empathy including exclusively neurons with mirror-like properties. Indeed, we report that single neurons in AI encode empathy, especially concerning the aversive experience, but not with a mirror-like pattern as the dominant profile.

The clinical relevance of empathy studies

Our ability to respond to other's distress empathically is a cornerstone for a successful social life. Paradoxically, this important interpersonal skill might also confer risk for depression and anxiety when present at extreme levels and in combination with certain individual characteristics or within particular contexts. Indeed, empathic reactions to others' distress that is excessively aversive and involves excessive cognitive perspective

taking may facilitate the emergence of internalizing problems, such as personal distress and excessive interpersonal guilt.

In common belief, empathy with other individuals is linked with positive interpersonal and intrapsychic outcomes, including better relationships with friends and partners (Chow et al., 2013; Cramer et al., 2010), increased social engagement (Bailey et al., 2008) and resilience (Shiner et al., 2012). However, several works showed associations between an impaired empathic capacity and psychopathological conditions such as conduct disorder and autism spectrum disorders (Miller et al., 1988; Decety et al., 2010; Bons et al., 2013). Few works have described potential psychopathological correlates of excessive empathy. However, as Zahn-Waxler and colleagues (1991) have suggested a model focused explicitly on girls, extreme empathy or a combination of enhanced sensitivity to others' distress and inadequate skills for coping with that distress may increase vulnerability to internalizing conditions marked by negative affect. Anyways, Tone and Tully (2014) sustain that also males who show excessive empathy are at high risk for internalizing conditions. Thus, the authors do not postulate distinct empathy-related pathways to internalizing problems for males and females.

Personal distress and Interpersonal guilt are the two main problems deriving from internalization. Personal distress is a maladaptive affective response to negative emotions in others. Interpersonal guilt, in contrast, is a maladaptive form of cognitive empathy that is driven by excessive and irrational altruistic concerns, such as unreasonable beliefs that one is responsible for alleviating the suffering for others and intense worries about harming others (O'Connor et al., 2007, 2002; Oakley et al., 2012; Zahn-Waxler et al., 2012).

Thus, the study of empathy has a considerable clinical relevance since atypical levels of empathy can lead to sociopathic behavior on one extreme and to stress and depression on the other.

Conclusions and perspectives

In conclusion, this work gives new insights in what has been known till now about the functionality of the anterior insula in primates. Because of its neuroanatomical position, this brain region has always been quite challenging to study, especially with invasive techniques such as electrophysiology. Indeed, most of the studies conducted in humans have used non invasive techniques such as fMRI. This insular cortex merits to be studied since it is involved in processing afferent bodily signals with strong connections to the limbic area. In humans, it is activated by felt and observed pain, and disgust. It is also activated by events like social rejection and grief, suggesting a more general role coding of the subjective quality of emotional experiences. Moreover, excessive empathy has been shown to lead to sociopathic behavior such as anxiety and depression.

Because of its biological similarities with humans, the non-human primate model has established itself as a gold standard over the past years. However, the difficulties and challenges in working with this species in laboratories are not to be underestimated. Especially in our research work, we trained the monkeys to perform a behavioral task that engaged higher cognitive functions, where they had to learn several associations for own and other's cues, and we paired this already difficult behavioral study with electrophysiological recordings in a deep brain area.

Bibliography

- Adolphs R (2009) The Social Brain: Neural Basis of Social Knowledge. *Annu Rev Psychol* 60:693–716.
- Amodio DM, Frith CD (2006) Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
- Atsak P, Orre M, Bakker P, Cerliani L, Roozendaal B, Gazzola V, Moita M, Keysers C (2011) Experience Modulates Vicarious Freezing in Rats: A Model for Empathy Ferrari PF, ed. *PLoS ONE* 6:e21855.
- Azzi JCB, Sirigu A, Duhamel J-R (2012) Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc Natl Acad Sci U S A* 109:2126–2131.
- Babineau BA, Bliss-Moreau E, Machado CJ, Toscano JE, Mason WA, Amaral DG (2011) Context-specific social behavior is altered by orbitofrontal cortex lesions in adult rhesus macaques. *Neuroscience* 179:80–93.
- Baez-Mendoza R, Harris CJ, Schultz W (2013) Activity of striatal neurons reflects social action and own reward. *Proc Natl Acad Sci* 110:16634–16639.
- Báez-Mendoza R, van Coeverden CR, Schultz W (2016) A neuronal reward inequity signal in primate striatum. *J Neurophysiol* 115:68–79.
- Ballesta S, Duhamel J-R (2015) Rudimentary empathy in macaques' social decision-making. *Proc Natl Acad Sci* 112:15516–15521.
- Bartal IB-A, Decety J, Mason P (2011) Empathy and Pro-Social Behavior in Rats. *Science* 334:1427–1430.
- Behrens TE., Hunt LT, Rushworth MF. (2009) The computation of social behavior. *Science* 324:1160.
- Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS (2008) Associative learning of social value. *Nature* 456:245–249.
- Bevacqua S, Cerasti E, Falcone R, Cervelloni M, Brunamonti E, Ferraina S, Genovesio A (2013) Macaque Monkeys Can Learn Token Values from Human Models through Vicarious Reward Chapouthier G, ed. *PLoS ONE* 8:e59961.
- Bird G, Viding E (2014) The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neurosci Biobehav Rev* 47:520–532.
- Blakemore SJ (2008) The social brain in adolescence. *Nat Rev Neurosci* 9:267–277.

- Bredy TW, Barad M (2008) Social modulation of associative fear learning by pheromone communication. *Learn Mem* 16:12–18.
- Brosnan SF (2013) Justice-and fairness-related behaviors in nonhuman primates. *Proc Natl Acad Sci* 110:10416–10423.
- Burkett JP, Andari E, Johnson ZV, Curry DC, de Waal FBM, Young LJ (2016) Oxytocin-dependent consolation behavior in rodents. *Science* 351:375–378.
- Caruana F, Jezzini A, Sbriscia-Fioretti B, Rizzolatti G, Gallese V (2011) Emotional and Social Behaviors Elicited by Electrical Stimulation of the Insula in the Macaque Monkey. *Curr Biol* 21:195–199.
- Catenoix H, Isnard J, Guénot M, Petit J, Remy C, Mauguière F (2008) The role of the anterior insular cortex in ictal vomiting: A stereotactic electroencephalography study. *Epilepsy Behav* 13:560–563.
- Chang SWC, Fagan NA, Toda K, Utevsky AV, Pearson JM, Platt ML (2015) Neural mechanisms of social decision-making in the primate amygdala. *Proc Natl Acad Sci* 112:16012–16017.
- Chang SWC, Gariépy J-F, Platt ML (2012) Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci* 16:243–250.
- Church, R. M. (1959). Emotional reactions of rats to the pain of others. *Journal of Comparative and Physiological Psychology*, 52 (2), 132–134.
- Cromwell HC (2003) Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in Primate Striatum. *J Neurophysiol* 89:2823–2838.
- de Vignemont F, Singer T (2006) The empathic brain: how, when and why? *Trends Cogn Sci* 10:435–441.
- Di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G (1992) Understanding motor events: a neurophysiological study. *Exp Brain Res* 91:176–180.
- Dunbar RI (1998) The social brain hypothesis. *brain* 9:10.
- Dunbar RIM, Shultz S (2007) Understanding primate brain evolution. *Philos Trans R Soc Lond B Biol Sci* 362:649–658.
- Falcone R, Brunamonti E, Ferraina S, Genovesio A (2012a) Monkeys Monitor Human Goals in a Nonmatch-to-Goal Interactive Task de Polavieja GG, ed. *PLoS ONE* 7:e32209.
- Falcone R, Brunamonti E, Genovesio A (2012b) Vicarious learning from human models in monkeys. *PloS One* 7:e40283.

- Flombaum JI, Santos LR (2005) Rhesus monkeys attribute perceptions to others. *Curr Biol CB* 15:447–452.
- Frith C, Frith U (2010) Learning from Others: Introduction to the Special Review Series on Social Neuroscience. *Neuron* 65:739–743.
- Frith CD, Frith U (2006) The Neural Basis of Mentalizing. *Neuron* 50:531–534.
- Frith, C.D., Frith, U., 1999. Interacting Minds - A Biological Basis. *Science* (80). 286, 1692–1695.
- Frith, U., Leslie, A.M., Morton, J., 1991. The cognitive basis of a biological disorder: autism. *Trends Neurosci.* 14(10), 433-438.
- Gonzalez-Liencre C, Juckel G, Tas C, Friebe A, Brüne M (2014) Emotional contagion in mice: The role of familiarity. *Behav Brain Res* 263:16–21.
- Gothard KM, Battaglia FP, Erickson CA, Spitler KM, Amaral DG (2007) Neural responses to facial expression and face identity in the monkey amygdala. *J Neurophysiol* 97:1671–1683.
- Guzman, Y. F., Tronson, N. C., Guedea, A., Huh, K. H., Gao, C., & Radulovic, J. (2009). Social modeling of conditioned fear in mice by non-fearful conspecifics. *Behav Brain Res*, 201 (1), 173–178.
- Hari R, Henriksson L, Malinen S, Parkkonen L (2015) Centrality of Social Interaction in Human Brain Function. *Neuron* 88:181–193.
- Haroush K, Williams ZM (2015) Neuronal Prediction of Opponent’s Behavior during Cooperative Social Interchange in Primates. *Cell* 160:1233–1245.
- Hosokawa T, Watanabe M (2012) Prefrontal Neurons Represent Winning and Losing during Competitive Video Shooting Games between Monkeys. *J Neurosci* 32:7662–7671.
- Jackson PL, Meltzoff AN, Decety J (2005) How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage* 24:771–779.
- Jeon D, Kim S, Chetana M, Jo D, Ruley HE, Lin S-Y, Rabah D, Kinet J-P, Shin H-S (2010) Observational fear learning involves affective pain system and Cav1.2 Ca²⁺ channels in ACC. *Nat Neurosci* 13:482–488.
- Jezzini A, Caruana F, Stoianov I, Gallese V, Rizzolatti G (2012) Functional organization of the insula and inner perisylvian regions. *Proc Natl Acad Sci* 109:10077–10082.

- Kim EJ, Kim ES, Covey E, Kim JJ (2010) Social Transmission of Fear in Rats: The Role of 22-kHz Ultrasonic Distress Vocalization Chapouthier G, ed. *PLoS ONE* 5:e15077.
- Knapska E, Mikosz M, Werka T, Maren S (2010) Social modulation of learning in rats. *Learn Mem* 17:35–42.
- Krolak-Salmon P, Hénaff M-A, Isnard J, Tallon-Baudry C, Guénot M, Vighetto A, Bertrand O, Mauguier F (2003) An attention modulated response to disgust in human ventral anterior insula. *Ann Neurol* 53:446–453.
- Lamm C, Batson CD, Decety J (2007) The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J Cogn Neurosci* 19:42–58.
- Langford, D. J., Crager, S. E., Shehzad, Z., Smith, S. B., Sotocinal, S. G., Levenstadt, J. S., et al. (2006). Social modulation of pain as evidence for empathy in mice. *Science*, 312 (5782), 1967–1970.
- Lau B, Salzman CD (2009) The rhythms of learning. *Nat Neurosci* 12:675–676.
- Li, Z., Lu, Y. F., Li, C. L., Wang, Y., Sun, W., He, T., et al. (2014). Social interaction with a cagemate in pain facilitates subsequent spinal nociception via activation of the medial prefrontal cortex in rats. *Pain*, 155 (7), 1253–1261.
- Lieberman MD (2007) Social Cognitive Neuroscience: A Review of Core Processes. *Annu Rev Psychol* 58:259–289.
- Lockwood PL, Apps MAJ, Roiser JP, Viding E (2015) Encoding of Vicarious Reward Prediction in Anterior Cingulate Cortex and Relationship with Trait Empathy. *J Neurosci* 35:13720–13727.
- Machado CJ, Bachevalier J (2006) The impact of selective amygdala, orbital frontal cortex, or hippocampal formation lesions on established social relationships in rhesus monkeys. *Behav Neurosci* 120:761.
- Martin LJ, Hathaway G, Isbester K, Mirali S, Acland EL, Niederstrasser N, Slepian PM, Trost Z, Bartz JA, Sapolsky RM, Sternberg WF, Levitin DJ, Mogil JS (2015) Reducing Social Stress Elicits Emotional Contagion of Pain in Mouse and Human Strangers. *Curr Biol* 25:326–332.
- Massen JJM, van den Berg LM, Spruijt BM, Sterck EHM (2010) Generous leaders and selfish underdogs: pro-sociality in despotic macaques. *PloS One* 5:e9734.
- MASSERMAN JH, WECHKIN S, TERRIS W (1964) “ALTRUISTIC” BEHAVIOR IN RHESUS MONKEYS. *Am J Psychiatry* 121:584–585.
- Morrison SE, Salzman CD (2009) The convergence of information about rewarding and aversive stimuli in single neurons. *J Neurosci Off J Soc Neurosci* 29:11471–11483.

- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Panksepp JB, Lahvis GP (2011) Rodent empathy and affective neuroscience. *Neurosci Biobehav Rev* 35:1864–1875.
- Paus T (2001) Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci* 2 Available at: <http://search.proquest.com/openview/241643074aa9caba173a2910f5ef1797/1?pq-origsite=gscholar&cbl=44265> [Accessed October 15, 2017].
- Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space and value. *Nat Neurosci* 16:340–348.
- Premack D., Woodruff G., 1978. *Behav. Brain Sci.* 4, 515.
- Rizzolatti G, Craighero L (2004) THE MIRROR-NEURON SYSTEM. *Annu Rev Neurosci* 27:169–192.
- Rizzolatti G, Sinigaglia C (2010) The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci* 11:264–274.
- Rolls ET (2000) The orbitofrontal cortex and reward. *Cereb Cortex N Y N* 1991 10:284–294.
- Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MFS (2006) Separate neural pathways process different decision costs. *Nat Neurosci* 9:1161–1168.
- Rushworth MF, Mars RB, Sallet J (2013) Are there specialized circuits for social cognition and are they unique to humans? *Curr Opin Neurobiol* Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0959438812001870> [Accessed May 3, 2013].
- Sallet J, Mars RB, Noonan MP, Andersson JL, O'Reilly JX, Jbabdi S, Crosson PL, Jenkinson M, Miller KL, Rushworth MFS (2011) Social Network Size Affects Neural Circuits in Macaques. *Science* 334:697–700.
- Sanders J, Mayford M, Jeste D (2013) Empathic Fear Responses in Mice Are Triggered by Recognition of a Shared Experience Lu L, ed. *PLoS ONE* 8:e74609.
- Santos LR, Nissen AG, Ferrugia JA (2006) Rhesus monkeys, *Macaca mulatta*, know what others can and cannot hear. *Anim Behav* 71:1175–1181.
- Schultz W (2015) Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* 95:853–951.
- Singer T (2004) Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science* 303:1157–1162.

Singer T (2006) The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci Biobehav Rev* 30:855–863.

Singer T, Lamm C (2009) The Social Neuroscience of Empathy. *Ann N Y Acad Sci* 1156:81–96.

Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.

Watanabe, S. (2015). Common experience modifies the reinforcing properties of methamphetamine injected cage mates but not morphine-injected cage mates in C57 mice. *Behavioural Pharmacology*, 26 (7 Spec No), 636–641.

Wise SP (2008) Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci* 31:599–608.

Yoshida K, Saito N, Iriki A, Isoda M (2011) Representation of Others' Action by Neurons in Monkey Medial Frontal Cortex. *Curr Biol*.

Yoshida K, Saito N, Iriki A, Isoda M (2012) Social error monitoring in macaque frontal cortex. *Nat Neurosci* 15:1307–1312.

Yusufishaq S, Rosenkranz JA (2013) Post-weaning social isolation impairs observational fear conditioning. *Behav Brain Res* 242:142–149.